

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:18:43 ; Search time 52 Seconds
(without alignments)
135.840 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	25	AAW42930	Immunogen
2	129	100.0	26	AAW42930	Synthetic
3	129	100.0	366	AAW42930	Sequence
4	129	100.0	854	AAW42930	Hepatitis
5	129	100.0	993	AAW42930	Sequence
6	129	100.0	993	AAW42930	Sequence
7	129	100.0	1077	AAW42930	Partial
8	129	100.0	1091	AAW42930	Translate
9	129	100.0	2227	AAW42930	Attenuate
10	129	100.0	2227	AAW42930	Hepatitis
11	129	100.0	2227	AAW42930	Amino aci
12	129	100.0	2227	AAW42930	Hepatitis
13	129	100.0	2227	AAW42930	Hepatitis
14	129	100.0	2227	AAW42930	Attenuate
15	129	100.0	2227	AAW42930	Wild-type
16	129	100.0	2227	AAW42930	Attenuat
17	129	100.0	2227	AAW42930	Wild type
18	129	100.0	2227	AAW42930	Hepatitis
19	124	96.1	2227	AAW42930	Sequence
20	124	96.1	2227	AAW42930	Amino aci
21	124	96.1	2227	AAW42930	Hepatitis
22	124	96.1	2227	AAW42930	Attenuat
23	117	90.7	839	AAW42930	Capsid re
24	103	79.8	20	AAW42930	Immunogen
25	103	79.8	21	AAW42930	Synthetic

26	65	50.4	20	2	AAW42925	Immunogen
27	65	50.4	21	4	AAW42925	Synthetic
28	51	39.5	150	4	AAW42925	Drosophil
29	49	38.0	420	4	AAW42925	Drosophil
30	49	38.0	444	2	AAW42925	H. pylori
31	49	38.0	686	6	AAW42925	Acinetoba
32	48.5	37.6	2000	6	AAW42925	Protein s
33	48	37.2	187	3	AAW42925	Eucalyptu
34	48	37.2	221	4	AAW42925	Drosophil
35	48	37.2	523	7	AAW42925	Enterohae
36	48	37.2	1176	4	AAW42925	Drosophil
37	47.5	36.8	86	3	AAW42925	Human sec
38	47.5	36.8	241	7	AAW42925	Human pro
39	47.5	36.8	468	7	AAW42925	Human pro
40	47.5	36.8	530	3	AAW42925	Human ORF
41	47.5	36.8	530	5	AAW42925	Human pol
42	47.5	36.8	530	5	AAW42925	Human gly
43	46.5	36.0	23	2	AAW42925	BMP-8 pep
44	46.5	36.0	137	2	AAW42925	Murine os
45	46.5	36.0	139	2	AAW42925	Mature mu

ALIGNMENTS

RESULT 1
AAW42930
ID AAW42930 standard; peptide; 25 AA.
XX
AC AAW42930;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1665.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-Al.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006891.
XX
PR 19-APR-1996; 96US-0015644P.
XX
PS (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 1997-S35831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. Compositions
CC containing the peptides can be used to induce an immune response to HAV
CC in a mammal. The peptides can also be used to detect the presence of
CC antibodies against HAV in mammalian serum. The peptides can also be used
CC to make an antibody against HAV by administering the peptide to a mammal
XX
SQ Sequence 25 AA;

Query Match 100.0%; Score 129; DB 2; Length 25;

ID	AAP50230	standard; protein; 366 AA.
XX		
AC	AAP50230;	
XX		
DT	28-NOV-1991	(first entry)
XX		
DE	Sequence of hepatitis A virus (HAV) surface protein (VP-1).	
XX		
KW	Hepatitis A virus vaccine; immunisation; monoclonal antibody;	
KW	diagnostic assay.	
XX		
OS	Hepatitis A virus.	
XX		
PN	EP138704-A.	
XX		
PD	24-APR-1985.	
XX		
PF	09-OCT-1984;	84EP-00402025.
XX		
PR	14-OCT-1983;	83US-00541836.
PR	02-MAR-1984;	84US-00585942.
XX		
PA	(MERI) MERCK & CO INC.	
XX		
PI	Hughes JV, Scolnick EM, Tomassini JE;	
XX		
DR	WFI; 1985-100818/17.	
DR	N-PSDB; AANS0274.	
XX		
PT	New hepatitis A virus surface protein - useful for binding to	
PT	neutralising antibodies to the virus.	
XX		
PS	Claim 21; Page 46-48; 49pp; English.	
XX		
CC	VP1 is isolated by solubilisation of the intact virus in an aq. anionic	
CC	surfactant and a reducing agent. The viral proteins are sepd. and the	
CC	protein of molecular wt. 33000 daltons is sepd	
XX		
SQ	Sequence 366 AA;	
	Query Match	100.0%; Score 129; DB 1; Length 366;
	Best Local Similarity	100.0%; Pred. No. 2.4e-11;
	Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps
QY	1	QRLKYAQEELSNEVLPPPRKMKGLF 25
Db	319	QRLKYAQEELSNEVLPPPRKMKGLF 343
RESULT 4		
AAP50287		
ID	AAP50287	standard; protein; 854 AA.
XX		
AC	AAP50287;	
XX		
DT	25-MAR-2003	(revised)
DT	30-NOV-1991	(first entry)
XX		
DE	Hepatitis A virus (HAV) peptide corresponding to the capsid protein	
DE	region of poliovirus RNA.	
XX		
KW	Hepatitis A virus assay; antigen; antibody.	
XX		
OS	Hepatitis A virus.	
XX		
PN	WO8501517-A.	
XX		
PD	11-APR-1985.	
XX		
PF	27-SEP-1984;	84WO-US001552.
XX		
PR	30-SEP-1983;	83US-00537911.
XX		

(MASI) MASSACHUSETTS INST TECHNOLOGY.
 PA Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 PI Racaniello VR;
 XX WPI; 1985-098846/16.
 DR N-PSDB; AAN50330.
 XX
 XX New hepatitis A virus CDNA - useful in assays for the virus and for
 PT prodn. of the viral antigen and antibodies to it.
 XX
 XX Example; Fig 7; 60pp; English.
 XX
 XX The inventors claim HAV cDNA and a method for producing it, whereby large
 CC ams. can be obt'd. economically. The cDNA is useful in the assay for
 CC detection of HAV quickly and easily and with high sensitivity and
 CC specificity. The HAV cDNA is also used in the prodn. of HAV antigen or
 CC antibodies to it. The antibodies may be monoclonal. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX
 XX Sequence 854 AA;
 SQ
 Query Match 100.0%; Score 129; DB 1; Length 854;
 Best Local Similarity 100.0%; Pred. No. 5.9e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMGGLF 834
 RESULT 5
 AAP50116
 ID AAP50116 standard; protein; 993 AA.
 XX
 AC AAP50116;
 XX
 DT 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 30-SEP-1991 (first entry)
 XX
 XX Sequence of Hepatitis A virus (HAV) immunogenic peptides VP-1, VP-2, VP-3
 DE and VP-4.
 XX
 XX Antigenic protein; immunogen; vaccine.
 XX
 XX Hepatitis A virus; (strain CR326).
 XX
 XX EP154587-A.
 XX
 XX 11-SEP-1985.
 XX
 XX 27-FEB-1985; 85EP-00400369.
 XX
 XX 02-MAR-1984; 84US-00585818.
 XX
 XX (MERI) MERCK & CO INC.
 XX
 XX Linemeyer DL, Menke JG, Rueben RG, Mitra SW;
 XX
 XX WPI; 1985-224964/37.
 DR N-PSDB; AAN50139.
 XX
 XX New nucleotide sequences coding for hepatitis A virus antigens - useful
 PT for eliciting normal immune response and in vaccines for protecting
 PT against the virus.
 XX
 XX Example; Page 11-17; 32pp; English.
 PS
 XX Within the sequence in AAN50139 is encoded the information necessary to
 CC make the antigenic proteins of HAV. The sequences encoding for the
 CC structural proteins begin at base 403. The key sub-unit sequences within
 CC VP-1, designated Sequences I,II,III,IV, and V, start, respectively at

CC 1982, 1963, 1999, 2146, 2347. Other nucleotide sequences which are
 CC valuable as encoding antigenic proteins are the sequences from base 1749
 CC to base 2722; from base 1987 to base 2980 and from base 1644 to base
 CC 2722. The sequence from base 1749 to base 2722 is esp. valuable as a
 CC vector for producing antigen protein. Sequences II-V are claimed. X in
 CC AAP50116 denotes the translation of a stop codon. (Updated on 25-MAR-2003
 CC to correct PI field.) (Updated on 24-OCT-2003 to standardise OS field)
 XX
 XX Sequence 993 AA;
 SQ
 Query Match 100.0%; Score 129; DB 1; Length 993;
 Best Local Similarity 100.0%; Pred. No. 7e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
 DB 946 QRLKYAQEELSNEVLPPPRKMGGLF 970
 RESULT 6
 AAP50231
 ID AAP50231 standard; protein; 993 AA.
 XX
 AC AAP50231;
 XX
 DT 28-NOV-1991 (first entry)
 XX
 DE Sequence encoded by partial sequence of hepatitis A virus (HAV),
 DE including surface protein (VP-1).
 XX
 XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.
 XX
 XX Hepatitis A virus.
 OS
 XX
 FH Key Location/Qualifiers
 FT Protein 628..993
 FT /note= "claimed; X denotes translated stop codons and
 FT unspecified triplets"
 XX
 PN EP138704-A.
 XX
 XX 24-APR-1985.
 XX
 XX 09-OCT-1984; 84EP-00402025.
 XX
 XX 14-OCT-1983; 83US-00541836.
 PR 02-MAR-1984; 84US-00585942.
 XX
 XX (MERI) MERCK & CO INC.
 XX
 XX Hughes JV, Scolnick EM, Tomassini JE;
 PI
 XX WPI; 1985-100818/17.
 DR N-PSDB; AAN50274.
 XX
 XX New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus.
 XX
 XX Disclosure; Page 17-23; 49pp; English.
 PS
 XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
 CC surfactant and a reducing agent. The viral proteins are sepd. and the
 CC protein of molecular wt. 33000 daltons is sepd
 XX
 XX Sequence 993 AA;
 SQ
 Query Match 100.0%; Score 129; DB 1; Length 993;
 Best Local Similarity 100.0%; Pred. No. 7e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
 |||||

Db 946 QRLKYAQEELSNEVLPPPRKMKGLF 970

RESULT 7

AAW95559
ID AAW95559 standard; protein; 1077 AA.

XX AC AAW95559;
XX DT 28-APR-1999 (first entry)

XX DE A partial hepatitis A virus (HAV) protein.
XX KW Hepatitis A virus protein; HAV; P2 region;
XX cell-culture-adapted HAV strain; infection; accelerated growth.

XX OS Hepatitis A virus.
XX PN US5849562-A.
XX PD 15-DEC-1998.
XX PF 06-JUN-1995; 95US-00468926.

XX PR 30-SEP-1983; 83US-00537911.
XX PR 27-SEP-1984; 84US-00654942.
XX PR 06-OCT-1988; 88US-00256135.
XX PR 06-NOV-1991; 91US-00788262.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Emerson SU, Purcell RH;
XX DR WPI; 1999-094412/08.
XX DR N-PSDB; AAX01006.

XX PT Chimeric hepatitis A virus strains - with P2 region from cell-culture-adapted strain in wild-type genome.
XX PS Disclosure; Fig 7A-L; 36pp; English.
XX CC The present sequence represents a partial hepatitis A virus (HAV) protein. The specification describes a DNA construct consisting of a wild-type HAV genome in which the P2 region is replaced by the P2 region from a cell-culture-adapted HAV strain. The construct is used to demonstrate that mutations in the P2 region of a cell-culture-adapted HAV strain are sufficient for establishment of infection and accelerated growth in cell culture

XX SQ Sequence 1077 AA;
Query Match 100.0%; Score 129; DB 2; Length 1077;
Best Local Similarity 100.0%; Pred. No. 7.6e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 1033 QRLKYAQEELSNEVLPPPRKMKGLF 1057

RESULT 8

AAW92426
ID AAW92426 standard; protein; 1091 AA.

XX AC AAR32426;
XX DT 27-AUG-2003 (revised)

XX DT 25-MAR-2003 (revised)
XX DT 17-DEC-2001 (revised)
XX DT 10-JUN-1993 (first entry)

XX DE Translated from 5' region of Hepatitis A Virus genomic clone.
XX DT 24-OCT-2003 (revised)
XX DT 27-AUG-2003 (revised)

KW HAV HM-175; chronic liver disease; picornavirus.
XX OS Hepatitis A virus.
XX FH Key
FT Region Location/Qualifiers
FT 1..711
FT /note= "X's correspond to nonsense codons, i.e. this region is not an ORF"
FT 238..1091
FT /label= ORF
FT /note= "second putative initiation codon at position 240"
XX PN USN7788262-N.
XX PD 15-DEC-1992.
XX PF 06-NOV-1991; 91US-00788262.
XX PR 30-SEP-1983; 88US-00536911.
XX PR 27-SEP-1984; 84US-00654942.
XX PR 06-OCT-1988; 88US-00256135.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX PI Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
XX PI Racaniello VR, Baroudy BM, Emerson SU;
XX DR WPI; 1993-067429/08.
XX DR N-PSDB; AAQ36934.

XX PT Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn. of antigen and antibodies.
XX PS Disclosure; Fig 7; 65pp; English.

XX CC HAV virion RNA was extracted from the livers of marmosets which had been inoculated with HAV (the HAV had previously been passaged twice in marmosets). The RNA was used to prepare ds cDNA clones by standard methods. Clones contg. inserts which hybridised to RNA from HAV-infected African Green Monkey Kidney cells were selected for further analysis. A 7.4kb restriction map (about 99% of the HAV genome) was constructed from the 5'-terminus was determined. The sequence of the first 3.3kb (approx.) from the entire clone and an open reading frame was identified starting at position 238. A comparison of the predicted HAV amino acid sequences with the known capsid protein sequences of other picornaviruses (poliovirus, foot and mouth disease virus and encephalomyelitis virus) revealed areas of local homology. (Note: Revised entry submitted to correct the patent number format of US Government-owned NTIS applications to prevent clashes with ongoing US granted patent numbers. For further information please visit the Derwent web site at www.derwent.com/dwpi/updates/ntis.us.html.) (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 1091 AA;
Query Match 100.0%; Score 129; DB 2; Length 1091;
Best Local Similarity 100.0%; Pred. No. 7.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 1047 QRLKYAQEELSNEVLPPPRKMKGLF 1071

RESULT 9

AAW05697
ID AAR05697 standard; protein; 2227 AA.

XX AC AAR05697;
XX DT 24-OCT-2003 (revised)

XX DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)
 DT 15-AUG-1990 (first entry)
 XX Attenuated hepatitis A virus.
 XX Hepatitis A virus; vaccine; attenuated.
 KW Hepatitis A virus; strain HM-175.
 XX
 OS Hepatitis A virus; Location/Qualifiers
 XX Key 1. .23
 FH Region /label= VP4 = 1A
 FT Region 24. .245
 FT Region /label= VP2 = 1B
 FT Region 246. .491
 FT Region /label= VP3 = 1C
 FT Region 492. .791
 FT Region /label= VP1 = 1D
 FT Region 792. .980
 FT Region /label= 2A
 FT Region 981. .1087
 FT Region /label= 2B
 FT Region 1088. .1422
 FT Region /label= 2C
 FT Region 1423. .1496
 FT Region /label= 3A
 FT Region 1497. .1519
 FT Region /label= 3B = VPg
 FT Region 1520. .1738
 FT Region /label= 3C
 FT Region 1739. .2227
 FT Region /label= 3D
 XX US4894228-A.
 PN
 XX
 XX 16-JAN-1990.
 PD
 XX
 XX 12-JUL-1988; 88US-00217824.
 PF
 XX
 XX 19-SEP-1984; 84US-00652067.
 PR
 XX 09-SEP-1986; 86US-00905146.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PA
 XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstone SM;
 PI Daemer RJ, Gust ID;
 PI
 XX WPI: 1990-075557/10.
 DR N-PSDB; AAQ03512.
 DR
 XX Vaccine against hepatitis A virus infection - comprises novel attenuated
 PT hepatitis A virus strain.
 PT
 XX Claim 1; Fig 1; 18pp; English.
 PS
 XX The attenuated HAV is useful for inducing protective immunity against
 CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
 CC several nucleotide changes distributed throughout the genome, is
 CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
 CC suitable for use as an HAV vaccine. It is noted that not all the changes
 CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
 CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
 CC standardise OS field)
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 129; DB 2; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMGILF 25
 |||||||||||||||||||||||||||||

Db 810 QRLKYAQEELSNEVLPPPRKMGILF 834
 RESULT 10
 AAW34074
 ID AAW34074 standard; protein; 2227 AA.
 XX
 AC AAW34074;
 XX
 XX 17-OCT-2003 (revised)
 DT 27-APR-1998 (first entry)
 DT
 XX Hepatitis A virus HM-175 protein sequence.
 DE
 XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;
 KW vaccine.
 KW
 XX Hepatitis A virus; HM-175.
 OS
 XX Key Location/Qualifiers
 FH Protein 1. .23
 FT /label= VP4
 FT Protein 24. .245
 FT /label= VP2
 FT Protein 246. .491
 FT /label= VP3
 FT Protein 492. .791
 FT /label= VP1
 FT Protein 792. .980
 FT /label= 2A
 FT Protein 981. .1087
 FT /label= 2B
 FT Protein 1088. .1422
 FT /label= 2C
 FT Protein 1423. .1496
 FT /label= 3A
 FT Protein 1497. .1519
 FT /label= 3B
 FT Protein 1520. .1738
 FT /label= 3C
 FT Protein 1739. .2227
 FT /label= 3D
 XX WO9740166-A2.
 PN
 XX 30-OCT-1997.
 PD
 XX
 XX 18-APR-1997; 97WO-US006506.
 PF
 XX
 XX 19-APR-1996; 96US-0015642P.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Raychaudhuri G, Emerson SU, Purcell RH;
 PI
 XX WPI: 1997-535850/49.
 DR N-PSDB; AAT93023.
 DR
 XX Human attenuated HAV genome containing simian HAV 2C gene - useful as
 PT vaccines against HAV infection.
 PT
 XX Disclosure; Fig 13A-D; 66pp; English.
 PS
 XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-
 CC 175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained
 CC by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA
 CC construct (1) comprises a genome of HAV, where the genome is a human
 CC attenuated HAV genome in which a region of the 2C gene has been replaced
 CC by a corresponding region from a 2C gene of a simian ACM-27 HAV genome
 CC (see AAT93024). The region of the 2C gene from ACM-27 contained in the
 CC construct preferably encodes amino acids 120-328 of the 2C protein, amino
 CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript
 CC of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3)

CC a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host
CC cell containing the HAV of (3). (1) or its RNA transcript, can be used as
CC a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can
CC also be used to stimulate the production of protective antibodies in the
CC mammal. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 2; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
|||||
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 11

AAE18609
ID AAE18609 standard; protein; 2227 AA.

XX AC AAE18609;

XX DT 15-JAN-2001 (first entry)

XX DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX HAV 4380.

XX OS Hepatitis A virus.

XX PN US6113912-A.

XX PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-00475886.

XX PR 18-SEP-1992; 92US-00947338.

XX PR 17-SEP-1993; 93WO-US0008610.

XX PR 17-APR-1995; 95US-00397232.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX DR WPI; 2000-586464/55.

XX DR N-PSDB; AAA75478.

XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type.

XX PS Disclosure; Col 93-104; 72pp; English.

XX The present sequence is derived from a live attenuated hepatitis A virus
XX (HAV) of the invention, designated HAV 4380. The sequence is produced by
XX modifying wild type HAV strain HM-174. The HAV of the invention are
XX adapted to growth in the human fibroblast-like cell line MRC-5. The HAV
XX is able to propagate in MRC-5 cells and retain appropriate attenuation.
XX It is useful as a live vaccine for prophylaxis of hepatitis A in humans
XX and other primates

SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
|||||
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 12
AAE18607
ID AAE18607 standard; protein; 2227 AA.

XX AC AAE18607;

XX DT 15-JAN-2001 (first entry)

XX DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

XX OS Hepatitis A virus.

XX PN US6113912-A.

XX PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-00475886.

XX PR 18-SEP-1992; 92US-00947338.

XX PR 17-SEP-1993; 93WO-US0008610.

XX PR 17-APR-1995; 95US-00397232.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX DR WPI; 2000-586464/55.

XX DR N-PSDB; AAA75476.

XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type.

XX PS Disclosure; Fig 6A-K; 72pp; English.

XX The present sequence is derived from a wild type hepatitis A virus (HAV)
XX strain HM-174. The sequence is modified to produce HAV which are adapted
XX to growth in the human fibroblast-like cell line MRC-5. The HAV is able
XX to propagate in MRC-5 cells and retain appropriate attenuation. It is
XX useful as a live vaccine for prophylaxis of hepatitis A in humans and
XX other primates

SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
|||||
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 13

AAE19899

ID AAE19899 standard; protein; 2227 AA.

XX AC AAE19899;

XX DT 18-JUN-2002 (first entry)

XX DE Hepatitis A virus (HAV) protein.

XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.

XX OS Hepatitis A virus.

XX PN WO200213855-A2.

XX

PD 21-FEB-2002.
 XX 15-AUG-2001; 2001WO-IB001808.
 PF 17-AUG-2000; 2000US-0225767P.
 PR 29-AUG-2000; 2000US-0229175P.
 PR 03-NOV-2000; 2000US-00705547.
 XX (TRIP-) TRIPEP AB.
 PA Sallberg M, Hultgren C;
 PI WPI; 2002-241837/29.
 XX N-PSDB; AAD31766.
 DR Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 XX Claim 11; Page 82-87; 120pp; English.
 PS The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to a
 CC hepatitis C antigen in humans, domestic, sport or pet species and as
 CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is hepatitis A virus (HAV) protein
 XX
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 129; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMGGLF 834
 RESULT 14
 ABG31729
 ID ABG31729 standard; protein; 2227 AA.
 XX AC ABG31729;
 XX 29-AUG-2003 (revised)
 DT 29-NOV-2002 (first entry)
 DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
 XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
 KW HAV 4380.
 XX Hepatitis A virus; strain HM-175.
 OS US6423318-B1.
 PN 23-JUL-2002.
 PD 31-AUG-2000; 2000US-00653499.
 XX 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 PR 07-JUN-1995; 95US-00475886.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 WPI; 2002-680946/73.
 XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 in MRC-5 cells, useful for preparing a vaccine against HAV infection.
 PS Disclosure; Fig 6; 71pp; English.
 CC The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus
 CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
 CC field)
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 129; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMGGLF 834
 RESULT 15
 ABG31727
 ID ABG31727 standard; protein; 2227 AA.
 XX AC ABG31727;
 XX 29-AUG-2003 (revised)
 DT 29-NOV-2002 (first entry)
 DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
 XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
 OS Hepatitis A virus; strain HM-175.
 PN US6423318-B1.
 PD 23-JUL-2002.
 PD 31-AUG-2000; 2000US-00653499.
 XX 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 PR 07-JUN-1995; 95US-00475886.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 WPI; 2002-680946/73.
 XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 in MRC-5 cells, useful for preparing a vaccine against HAV infection.
 PS Disclosure; Fig 6; 71pp; English.
 CC The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus
 CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
 CC field)
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 129; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR N-PSDB; ABS52789.
 XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
 XX Disclosure; Col 93-104; 71pp; English.
 XX The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents an attenuated
 CC hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to
 CC standardise OS field)
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 129; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMGGLF 834
 RESULT 15
 ABG31727
 ID ABG31727 standard; protein; 2227 AA.
 XX AC ABG31727;
 XX 29-AUG-2003 (revised)
 DT 29-NOV-2002 (first entry)
 DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
 XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
 OS Hepatitis A virus; strain HM-175.
 PN US6423318-B1.
 PD 23-JUL-2002.
 PD 31-AUG-2000; 2000US-00653499.
 XX 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 PR 07-JUN-1995; 95US-00475886.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 WPI; 2002-680946/73.
 XX N-PSDB; ABS52787.
 XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 in MRC-5 cells, useful for preparing a vaccine against HAV infection.
 PS Disclosure; Fig 6; 71pp; English.
 CC The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus
 CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
 CC field)
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 129; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:26:44 ; Search time 18 Seconds
(without alignments)
71.703 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	1091	6	Patent No. 5516630-2
2	129	100.0	2227	3	US-08-475-886-2
3	129	100.0	2227	3	US-08-475-886-6
4	129	100.0	2227	3	US-08-397-232-2
5	129	100.0	2227	3	US-08-397-232-4
6	129	100.0	2227	3	US-09-171-387-2
7	129	100.0	2227	4	US-09-653-499-2
8	129	100.0	2227	4	US-09-653-499-6
9	129	100.0	2227	4	US-10-104-966-12
10	129	100.0	2227	4	US-10-135-988-2
11	129	100.0	2227	4	US-10-135-988-6
12	124	96.1	2227	3	US-08-475-886-4
13	124	96.1	2227	4	US-09-653-499-4
14	124	96.1	2227	4	US-10-135-988-4
15	117	90.7	839	1	US-08-087-016-2
16	49	38.0	686	4	US-09-328-352-4303
17	47	36.4	180	4	US-09-252-991A-18339
18	46.5	36.0	23	1	US-07-800-364B-4
19	46.5	36.0	23	5	PCT-US91-03388-4
20	46.5	36.0	139	1	US-08-278-729A-8
21	46.5	36.0	139	1	US-08-155-343A-8
22	46.5	36.0	139	1	US-08-406-672-8
23	46.5	36.0	139	1	US-08-643-563A-8
24	46.5	36.0	139	1	US-08-643-763A-8
25	46.5	36.0	139	1	US-08-462-623-8
26	46.5	36.0	139	1	US-08-451-953A-8
27	46.5	36.0	139	2	US-08-445-468A-8

28	46.5	36.0	139	2	US-08-461-397A-8	Sequence 8, Appli
29	46.5	36.0	139	2	US-08-912-088-8	Sequence 8, Appli
30	46.5	36.0	139	3	US-08-278-730A-8	Sequence 8, Appli
31	46.5	36.0	139	3	US-08-445-467-8	Sequence 8, Appli
32	46.5	36.0	139	3	US-08-480-515A-8	Sequence 8, Appli
33	46.5	36.0	139	3	US-08-414-033A-8	Sequence 8, Appli
34	46.5	36.0	139	3	US-08-271-556A-6	Sequence 6, Appli
35	46.5	36.0	139	3	US-08-440-894A-8	Sequence 8, Appli
36	46.5	36.0	139	4	US-09-170-936-8	Sequence 8, Appli
37	46.5	36.0	139	4	US-08-461-113-8	Sequence 8, Appli
38	46.5	36.0	139	4	US-08-456-033-8	Sequence 7, Appli
39	46.5	36.0	139	4	US-08-643-321-7	Sequence 8, Appli
40	46.5	36.0	139	4	US-09-464-206-8	Sequence 8, Appli
41	46.5	36.0	139	5	US-08-404-113A-8	Sequence 8, Appli
42	46.5	36.0	139	5	PCT-US92-01968-8	Sequence 8, Appli
43	46.5	36.0	139	5	PCT-US93-07190-8	Sequence 8, Appli
44	46.5	36.0	139	5	PCT-US93-07231-8	Sequence 8, Appli
45	46.5	36.0	139	5	PCT-US93-08742-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
5516630-2
Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTEIN, STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS

NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/788,262
FILING DATE: 06-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 256,135
FILING DATE: 06-OCT-1988
APPLICATION NUMBER: 654,942
FILING DATE: 27-SEP-1984
APPLICATION NUMBER: 537,911
FILING DATE: 30-SEP-1983
SEQ ID NO:2
LENGTH: 1091

Query Match 100.0%; Score 129; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 7.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 1047 QRLKYAQEELSNEVLPPPRKMKGLF 1071

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 3
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 4
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 5
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 6
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
GENERAL INFORMATION:
APPLICANT: RAYCHAUDHURI, GOPA;
EMERSON, SUZANNE, U.;
PURCELL, ROBERT, H.
TITLE OF INVENTION: SIMIAN-HUMAN HAV
HAVING A CHIMERIC 2C PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997

```
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 7
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match      100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 8
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
```

```
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match      100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 9
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 10
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
```

```
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAQEELSNEVLPPPRKMKGLF 25
   |||||
Db 810 ORLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 11
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAQEELSNEVLPPPRKMKGLF 25
   |||||
Db 810 ORLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 12
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-4

Query Match          96.1%; Score 124; DB 4; Length 2227;
Best Local Similarity 96.0%; Pred. No. 1e-10;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAQEELSNEVLPPPRKMKGLF 25
   |||||
Db 810 ORLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 13
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match          96.1%; Score 124; DB 4; Length 2227;
Best Local Similarity 96.0%; Pred. No. 1e-10;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAQEELSNEVLPPPRKMKGLF 25
   |||||
Db 810 ORLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 14
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
```



```

; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match          96.1%; Score 124; DB 4; Length 2227;
Best Local Similarity 96.0%; Pred. No. 1e-10; Indels 0; Gaps 0;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 810 QRLKYAQEELSSEVLPPPRKMGGLF 834

RESULT 15
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135
; GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2

Query Match          90.7%; Score 117; DB 1; Length 839;
Best Local Similarity 84.0%; Pred. No. 4.4e-10;
Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 809 QRFKYARELSNEILPPPRKMGGLF 833
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:24:53 ; Search time 43 Seconds
(without alignments)
181.721 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLKVAQEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 1285345

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	129	100.0	25	10	US-09-171-432A-47
2	129	100.0	352	14	US-10-272-459-45
3	129	100.0	980	14	US-10-272-459-41
4	129	100.0	2227	9	US-09-929-955-12
5	129	100.0	2227	13	US-10-104-966-12
6	129	100.0	2227	13	US-10-135-988-2
7	129	100.0	2227	13	US-10-135-988-6
8	129	100.0	2227	16	US-10-719-619-12
9	125	96.9	836	14	US-10-272-459-40
10	124	96.1	2227	13	US-10-135-988-4
11	103	79.8	20	10	US-09-171-432A-41
12	65	50.4	20	10	US-09-171-432A-42
13	54.5	42.2	306	16	US-10-437-963-156882
14	52	40.3	1193	15	US-10-369-493-5713
15	51	39.5	105	16	US-10-437-963-163647

16	51	39.5	1240	15	US-10-369-493-4031	Sequence 4031, Ap
17	51	39.5	6642	15	US-10-369-493-5013	Sequence 5013, Ap
18	50	38.8	1168	15	US-10-369-493-3980	Sequence 3980, Ap
19	49	38.0	442	12	US-10-335-977-5788	Sequence 5788, Ap
20	49	38.0	444	12	US-10-335-977-5789	Sequence 5789, Ap
21	48	37.2	63	12	US-10-424-599-254867	Sequence 254867, Ap
22	48	37.2	300	15	US-10-369-493-10595	Sequence 10595, A
23	48	37.2	1222	16	US-10-437-963-151046	Sequence 151046, A
24	48	37.2	4473	16	US-10-437-963-201113	Sequence 201113, A
25	48	37.2	4737	16	US-10-437-963-201116	Sequence 201116, A
26	47.5	36.8	241	15	US-10-104-047-2501	Sequence 2501, Ap
27	47.5	36.8	468	15	US-10-104-047-3508	Sequence 3508, Ap
28	47.5	36.8	530	15	US-10-108-260A-4331	Sequence 4331, Ap
29	47	36.4	65	12	US-10-424-599-237201	Sequence 237201, A
30	47	36.4	259	12	US-10-424-599-234250	Sequence 234250, A
31	47	36.4	355	12	US-10-425-114-44229	Sequence 44229, A
32	47	36.4	592	15	US-10-369-493-1534	Sequence 1534, Ap
33	46.5	36.0	99	12	US-10-424-599-246519	Sequence 246519, A
34	46.5	36.0	139	8	US-08-260-675-8	Sequence 8, Appli
35	46.5	36.0	139	12	US-09-952-318A-8	Sequence 8, Appli
36	46.5	36.0	139	14	US-10-050-050-8	Sequence 0, Appli
37	46.5	36.0	139	16	US-10-385-064-8	Sequence 8, Appli
38	46.5	36.0	235	12	US-10-425-114-57289	Sequence 57289, A
39	46.5	36.0	399	8	US-08-957-425-27	Sequence 27, Appl
40	46.5	36.0	399	8	US-08-260-675-23	Sequence 23, Appl
41	46.5	36.0	399	12	US-09-952-318A-23	Sequence 23, Appl
42	46.5	36.0	399	12	US-10-321-799-27	Sequence 27, Appl
43	46.5	36.0	399	14	US-10-122-026-8	Sequence 8, Appli
44	46.5	36.0	399	14	US-10-050-050-23	Sequence 23, Appli
45	46.5	36.0	399	14	US-10-350-747-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-171-432A-47
; Sequence 47, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..25
; OTHER INFORMATION: /label= YK-1665
US-09-171-432A-47

Query Match      100.0%; Score 129; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

RESULT 2
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PFI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
US-10-272-459-45

Query Match      100.0%; Score 129; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 182 QRLKYAQEELSNEVLPPPRKMKGLF 206

RESULT 3
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PFI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 115.5 kDa
US-10-272-459-41
```

```
Query Match      100.0%; Score 129; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 6.4e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 4
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match      100.0%; Score 129; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 5
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      100.0%; Score 129; DB 13; Length 2227;
```

```

Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMGGLF 834

RESULT 6
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 129; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMGGLF 834

RESULT 7
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 129; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGGLF 25

```

```

Db 810 QRLKYAQEELSNEVLPPPRKMGGLF 834

RESULT 8
US-10-719-619-12
; Sequence 12, Application US/10719619
; Publication No. US20040086529A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1C
; CURRENT APPLICATION NUMBER: US/10/719,619
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 10/104,966
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-719-619-12

Query Match 100.0%; Score 129; DB 16; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMGGLF 834

RESULT 9
US-10-272-459-40
; Sequence 40, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 40
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 94 kDa
US-10-272-459-40

Query Match 96.9%; Score 125; DB 14; Length 836;
Best Local Similarity 96.0%; Pred. No. 2.1e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMGGLF 834

RESULT 10
US-10-135-988-4

```

```
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US 110-135-988-4
```

```
Query Match 96.1%; Score 124; DB 13; Length 2227;
Best Local Similarity 96.0%; Pred. No. 8.4e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QRLKYAQEELSNEVLPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPRKMKGLF 834
```

```
RESULT 11
US 09-171-432A-41
; Sequence 41, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1318
US-09-171-432A-41
Query Match 50.4%; Score 65; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 14 VLPPPRKMKGLF 25
DB 1 VLPPPRKMKGLF 12
```

```
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1317
US-09-171-432A-41
Query Match 79.8%; Score 103; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPRK 20
DB 1 QRLKYAQEELSNEVLPPRK 20
RESULT 12
US 09-171-432A-42
; Sequence 42, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1318
US-09-171-432A-42
Query Match 50.4%; Score 65; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 14 VLPPPRKMKGLF 25
DB 1 VLPPPRKMKGLF 12
```

```

RESULT 13
US-10-437-963-156882
; Sequence 156882, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156882
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MKT4530_56507C.1.pap
US-10-437-963-156882

Query Match          42.2%; Score 54.5; DB 16; Length 306;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 12; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

Qy 7 QEELSN---EVLPPPRKMKGL 24
Db 136 QRLNNAMREVLPPVDMKGI 156
      ||| ||||| |||
      :|| :||| :|||

RESULT 14
US-10-369-493-5713
; Sequence 5713, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5713
; LENGTH: 1193
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5713

Query Match          40.3%; Score 52; DB 15; Length 1193;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVL 16
Db 340 QRIKYAREILQKELLP 355
      ||| ||| ||| |||

RESULT 15
US-10-437-963-163647

```

```

; Sequence 163647, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163647
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MKT4530_62623C.1.pap
US-10-437-963-163647

Query Match          39.5%; Score 51; DB 16; Length 105;
Best Local Similarity 55.0%; Pred. No. 12;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RLKYAQEELSNEVLPPPRKM 21
Db 31 RVKPAWEKSEEDHPPPREM 50
      ||| ||| ||| ||| |||
      :|| :||| :||| :|||

Search completed: July 15, 2004, 15:28:27
Job time : 44 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:25:54 ; Search time 174 Seconds
(without alignments)
140.238 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLKIAQEELSNEVLPPRRMKGLF 25

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents AA Main:*
- 1: /cgn2_6/ptodata/2/paa/PCTUS COMB. pep.*
 - 2: /cgn2_6/ptodata/2/paa/US06 COMB. pep.*
 - 3: /cgn2_6/ptodata/2/paa/US07 COMB. pep.*
 - 4: /cgn2_6/ptodata/2/paa/US080 COMB. pep.*
 - 5: /cgn2_6/ptodata/2/paa/US081 COMB. pep.*
 - 6: /cgn2_6/ptodata/2/paa/US082 COMB. pep.*
 - 7: /cgn2_6/ptodata/2/paa/US083 COMB. pep.*
 - 8: /cgn2_6/ptodata/2/paa/US084 COMB. pep.*
 - 9: /cgn2_6/ptodata/2/paa/US085 COMB. pep.*
 - 10: /cgn2_6/ptodata/2/paa/US086 COMB. pep.*
 - 11: /cgn2_6/ptodata/2/paa/US087 COMB. pep.*
 - 12: /cgn2_6/ptodata/2/paa/US088 COMB. pep.*
 - 13: /cgn2_6/ptodata/2/paa/US089 COMB. pep.*
 - 14: /cgn2_6/ptodata/2/paa/US090 COMB. pep.*
 - 15: /cgn2_6/ptodata/2/paa/US091 COMB. pep.*
 - 16: /cgn2_6/ptodata/2/paa/US092 COMB. pep.*
 - 17: /cgn2_6/ptodata/2/paa/US093 COMB. pep.*
 - 18: /cgn2_6/ptodata/2/paa/US094 COMB. pep.*
 - 19: /cgn2_6/ptodata/2/paa/US095 COMB. pep.*
 - 20: /cgn2_6/ptodata/2/paa/US096 COMB. pep.*
 - 21: /cgn2_6/ptodata/2/paa/US097A COMB. pep.*
 - 22: /cgn2_6/ptodata/2/paa/US097B COMB. pep.*
 - 23: /cgn2_6/ptodata/2/paa/US098 COMB. pep.*
 - 24: /cgn2_6/ptodata/2/paa/US099A COMB. pep.*
 - 25: /cgn2_6/ptodata/2/paa/US099B COMB. pep.*
 - 26: /cgn2_6/ptodata/2/paa/US100 COMB. pep.*
 - 27: /cgn2_6/ptodata/2/paa/US101 COMB. pep.*
 - 28: /cgn2_6/ptodata/2/paa/US102 COMB. pep.*
 - 29: /cgn2_6/ptodata/2/paa/US103 COMB. pep.*
 - 30: /cgn2_6/ptodata/2/paa/US104 COMB. pep.*
 - 31: /cgn2_6/ptodata/2/paa/US106 COMB. pep.*
 - 32: /cgn2_6/ptodata/2/paa/US107 COMB. pep.*
 - 33: /cgn2_6/ptodata/2/paa/US60 COMB. pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description

1	129	100.0	25	15	US-09-171-432A-47	Sequence 47, Appl
2	129	100.0	26	26	US-10-031-088A-47	Sequence 47, Appl
3	129	100.0	352	1	PCT-US02-33121-45	Sequence 45, Appl
4	129	100.0	352	28	US-10-272-459-45	Sequence 41, Appl
5	129	100.0	980	1	PCT-US02-33121-41	Sequence 41, Appl
6	129	100.0	980	28	US-10-272-459-41	Sequence 2, Appl
7	129	100.0	2227	7	US-08-397-232-2	Sequence 2, Appl
8	129	100.0	2227	8	US-08-475-886-2	Sequence 12, Appl
9	129	100.0	2227	21	US-09-705-547-12	Sequence 12, Appl
10	129	100.0	2227	24	US-09-929-955-12	Sequence 12, Appl
11	129	100.0	2227	32	US-10-719-619-12	Sequence 12, Appl
12	129	100.0	2227	33	US-60-329-175-12	Sequence 12, Appl
13	125	96.9	836	1	PCT-US02-33121-40	Sequence 40, Appl
14	125	96.9	836	28	US-10-272-459-40	Sequence 40, Appl
15	124	96.1	2227	8	US-08-475-886-4	Sequence 4, Appl
16	117	90.7	839	8	US-08-446-883-2	Sequence 2, Appl
17	103	79.8	20	15	US-09-171-432A-41	Sequence 41, Appl
18	103	79.8	21	26	US-10-031-088A-41	Sequence 41, Appl
19	65	50.4	20	15	US-09-171-432A-42	Sequence 42, Appl
20	65	50.4	21	26	US-10-031-088A-42	Sequence 42, Appl
21	56.5	43.8	346	22	US-09-791-537-50866	Sequence 50866, A
22	54.5	42.2	306	30	US-10-437-963-156882	Sequence 156882, A
23	53	41.1	70	21	US-09-724-676-48970	Sequence 48970, A
24	53	41.1	70	21	US-09-724-676A-48970	Sequence 48970, A
25	53	41.1	99	22	US-09-758-470-416	Sequence 416, App
26	53	41.1	99	28	US-10-211-818-416	Sequence 416, App
27	53	41.1	123	21	US-09-724-676-48976	Sequence 48976, A
28	53	41.1	123	21	US-09-724-676-48982	Sequence 48982, A
29	53	41.1	123	21	US-09-724-676A-48976	Sequence 48976, A
30	53	41.1	123	21	US-09-724-676A-48982	Sequence 48982, A
31	53	41.1	137	22	US-09-758-470-397	Sequence 397, App
32	53	41.1	137	28	US-10-211-818-397	Sequence 397, App
33	53	41.1	144	22	US-09-760-476-2506	Sequence 2506, App
34	53	41.1	144	28	US-10-216-245-2506	Sequence 2506, App
35	53	41.1	146	22	US-09-758-470-645	Sequence 645, App
36	53	41.1	146	28	US-10-211-818-645	Sequence 645, App
37	53	41.1	221	21	US-09-724-676-48969	Sequence 48969, A
38	53	41.1	221	21	US-09-724-676A-48969	Sequence 48969, A
39	53	41.1	226	21	US-09-724-676-48968	Sequence 48968, A
40	53	41.1	226	21	US-09-724-676A-48968	Sequence 48968, A
41	53	41.1	274	21	US-09-724-676-48975	Sequence 48975, A
42	53	41.1	274	21	US-09-724-676-48981	Sequence 48981, A
43	53	41.1	274	21	US-09-724-676A-48975	Sequence 48975, A
44	53	41.1	274	21	US-09-724-676A-48981	Sequence 48981, A
45	53	41.1	279	21	US-09-724-676-48974	Sequence 48974, A

ALIGNMENTS

RESULT 1
US-09-171-432A-47
; Sequence 47, Application US/09171432A
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khuyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A

```
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..25
; OTHER INFORMATION: /label= YK-1665
;
US-09-171-432A-47

Query Match 100.0%; Score 129; DB 15; Length 25;
Best Local Similarity 100.0%; Pred. No. 9.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

RESULT 2
US-10-031-088A-47
; SEQUENCE 47, Application US/10031088A
; GENERAL INFORMATION:
; APPLICANT: Howard A. Fields
; TITLE OF INVENTION: Synthetic Peptides Immunoactive With Hepatitis A
; FILE OF INVENTION: Virus Antibodies
; FILE REFERENCE: 14114.0342PI
; CURRENT APPLICATION NUMBER: US/10/031.088A
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: PCT/US00/19267
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/144,412
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic HAV
; OTHER INFORMATION: P2A Peptide
US-10-031-088A-47

Query Match 100.0%; Score 129; DB 26; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

RESULT 3
PCT-US02-33121-45
; SEQUENCE 45, Application PC/TUS0233121
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
PCT-US02-33121-45

Query Match 100.0%; Score 129; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 182 QRLKYAQEELSNEVLPPPRKMKGLF 206

RESULT 4
US-10-272-459-45
; SEQUENCE 45, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match 100.0%; Score 129; DB 28; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 182 QRLKYAQEELSNEVLPPPRKMKGLF 206

RESULT 5
PCT-US02-33121-41
; SEQUENCE 41, Application PC/TUS0233121
; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
PCT-US02-33121-41
```

; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 115.5 kDa
PCT-US02-33121-41

Query Match 100.0%; Score 129; DB 1; Length 980;
Best Local Similarity 100.0%; Pred. No. 8.2e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 6

US-10-272-459-41
; Sequence 41, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match 100.0%; Score 129; DB 28; Length 980;
Best Local Similarity 100.0%; Pred. No. 8.2e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 7

US-08-397-232-2
; Sequence 2, Application US/08397232
; GENERAL INFORMATION:
; APPLICANT: Funkhouser, Ann W.
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; APPLICANT: D'Hondt, Eric
; TITLE OF INVENTION: Hepatitis A Virus Vaccines
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: PO Box 457 Spring House Corporate Cntr
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,232
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/947,338
; FILING DATE: 18-SEP-1992
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/US93/08610
; FILING DATE: 17-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: SBCP50110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-397-232-2

Query Match 100.0%; Score 129; DB 7; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 8

US-08-475-886-2
; Sequence 2, Application US/08475886
; GENERAL INFORMATION:
; APPLICANT: Funkhouser, Ann W.
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; APPLICANT: D'Hondt, Eric
; TITLE OF INVENTION: Hepatitis A Virus Vaccines
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: PO Box 457 Spring House Corporate Cntr
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,886
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/947,338
; FILING DATE: 18-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/397,232
; FILING DATE: 10-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: SBCP50110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-475-886-2

Query Match 100.0%; Score 129; DB 8; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
|||||
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 9
US-09-705-547-12
; Sequence 12, Application US/09705547
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.023AUS
; CURRENT APPLICATION NUMBER: US/09/705,547
; CURRENT FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-705-547-12

Query Match 100.0%; Score 129; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
|||||
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 10
US-09-929-955-12
; Sequence 12, Application US/09929955
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 129; DB 24; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834
|||||

RESULT 11
US-10-719-619-12
; Sequence 12, Application US/10719619
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1C
; CURRENT APPLICATION NUMBER: US/10/719,619
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 10/104,966
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-719-619-12

Query Match 100.0%; Score 129; DB 32; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
|||||
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 12
US-60-229-175-12
; Sequence 12, Application US/60229175
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.023PR
; CURRENT APPLICATION NUMBER: US/60/229,175
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-60-229-175-12

Query Match 100.0%; Score 129; DB 33; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
|||||
Db 810 QRLKYAQEELSNEVLPPPRKMGFLF 834

RESULT 13
PCT-US02-33121-40
; Sequence 40, Application PC/TUS0233121

```

; GENERAL INFORMATION:
; APPLICANT: CHIRON CORPORATION et al.
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI17955.003 / 2301-17955.40
; CURRENT APPLICATION NUMBER: PCT/US02/33121
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 94 kDa
PCT-US02-33121-40

Query Match          96.9%; Score 125; DB 1; Length 836;
Best Local Similarity 96.0%; Pred. No. 2.5e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 14
US-10-272-459-40
; Sequence 40, Application US/10272459
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 94 kDa
US-10-272-459-40

Query Match          96.9%; Score 125; DB 28; Length 836;
Best Local Similarity 96.0%; Pred. No. 2.5e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 15
US-08-475-886-4
; Sequence 4, Application US/08475886
; GENERAL INFORMATION:
; APPLICANT: Funkhouser, Ann W.
; APPLICANT: Emerson, Suzanne U.
; APPLICANT: Purcell, Robert H.
; APPLICANT: D'Hondt, Eric
; TITLE OF INVENTION: Hepatitis A Virus Vaccines
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: PO Box 457 Spring House Corporate Cntr
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA

```

```

; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,886
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/947,338
; FILING DATE: 18-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/397,232
; FILING DATE: 10-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: SBCP50110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-475-886-4

Query Match          96.1%; Score 124; DB 8; Length 2227;
Best Local Similarity 96.0%; Pred. No. 1.1e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

Search completed: July 15, 2004, 15:31:32
Job time : 174 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:24:28 ; Search time 19 Seconds
(without alignments)
97.946 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 395054 seqs, 74438689 residues

Total number of hits satisfying chosen parameters: 395054

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:*

- 1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	41.1	410	6	US-10-170-205E-3376
2	53	41.1	410	6	US-10-170-205E-3437
3	48	37.2	187	6	US-10-856-499-843
4	47.5	36.8	86	6	US-10-793-479-7640
5	47.5	36.8	416	6	US-10-170-205E-21306
6	47.5	36.8	538	6	US-10-170-205E-36793
7	47	36.4	129	6	US-10-767-701-40723
8	46	35.7	195	6	US-10-767-701-33120
9	46	35.7	224	7	US-60-579-902-9506
10	46	35.7	334	6	US-10-170-205E-14566
11	46	35.7	338	6	US-10-170-205E-14026
12	46	35.7	757	7	US-60-579-902-10359
13	45	34.9	155	6	US-10-767-701-40468
14	45	34.9	173	6	US-10-767-701-56206
15	45	34.9	185	6	US-10-170-205E-19914
16	45	34.9	208	6	US-10-170-205E-36334
17	45	34.9	221	6	US-10-170-205E-24785
18	45	34.9	223	6	US-10-784-480-84
19	45	34.9	374	7	US-60-556-841-1572
20	45	34.9	412	6	US-10-491-213-22
21	45	34.9	445	6	US-10-170-205E-25385
22	45	34.9	756	6	US-10-812-829-727
23	44.5	34.1	1897	6	US-10-170-205E-39346
24	44	34.1	919	5	US-09-248-796A-19408
25	43	33.3	213	6	US-10-100-683-9619
26	43	33.3	223	6	US-10-170-205E-35830

27	43	33.3	334	6	US-10-170-205E-25572	Sequence 25572, A
28	43	33.3	372	6	US-10-767-701-45758	Sequence 45758, A
29	43	33.3	408	5	US-09-248-796A-14380	Sequence 14380, A
30	43	33.3	531	6	US-10-170-205E-5204	Sequence 5204, Ap
31	43	33.3	782	7	US-60-556-841-3299	Sequence 3299, Ap
32	43	33.3	870	7	US-60-581-351-6927	Sequence 6927, Ap
33	43	33.3	916	1	PCT-US03-24982A-339	Sequence 339, App
34	43	33.3	2427	6	US-10-170-205E-17923	Sequence 17923, A
35	43	33.3	2696	6	US-10-170-205E-17922	Sequence 17922, A
36	42.5	32.9	112	6	US-10-856-499-844	Sequence 844, App
37	42.5	32.9	134	6	US-10-767-701-60747	Sequence 60747, A
38	42.5	32.9	139	6	US-10-767-701-33068	Sequence 33068, A
39	42.5	32.9	219	5	US-09-248-796A-18638	Sequence 18638, A
40	42	32.6	71	6	US-10-793-479-7020	Sequence 7020, Ap
41	42	32.6	93	6	US-10-170-205E-17446	Sequence 17446, A
42	42	32.6	122	1	PCT-US04-08833-4	Sequence 4, Appli
43	42	32.6	122	6	US-10-806-611-4	Sequence 4, Appli
44	42	32.6	125	1	PCT-US02-22858A-791	Sequence 791, App
45	42	32.6	125	1	PCT-US02-22858A-816	Sequence 816, App

ALIGNMENTS

RESULT 1

US-10-170-205E-3376
; Sequence 3376, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3376
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-3376

Query Match 41.1%; Score 53; DB 6; Length 410;
Best Local Similarity 50.0%; Pred. No. 7;
Matches 13; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

Qy 2 RLKYAQEEL--SNEVLPPPRKMKGLF 25
Db 58 RLQKAEELVKDNE LLPSPSLGSF 83

RESULT 2

US-10-170-205E-3437
; Sequence 3437, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3437
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-3437

Query Match 41.1%; Score 53; DB 6; Length 410;
Best Local Similarity 50.0%; Pred. No. 7;
Matches 13; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY 2 RLKYAOEEL--SNEVLPPPRKMKGLF 25
||: ||| ||: ||| |||
Db 58 RLKQAEELVKDNEVLPPPSLGSF 83

RESULT 3
US-10-856-499-843
; Sequence 843, Application US/10856499
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C2
; CURRENT APPLICATION NUMBER: US/10/856,499
; CURRENT FILING DATE: 2004-05-28
; NUMBER OF SEQ ID NOS: 2370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 843
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-856-499-843

Query Match 37.2%; Score 48; DB 6; Length 187;
Best Local Similarity 42.3%; Pred. No. 16;
Matches 11; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 2 RLKYAOEELSNEV--LPPPRKMKGLF 25
||: ||| ||: ||| |||
Db 116 RASQVRQELQAVDVKLPPRPQGNF 141

RESULT 4
US-10-793-479-7640
; Sequence 7640, Application US/10793479
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/10/793,479
; CURRENT FILING DATE: 2004-03-03
; PRIOR APPLICATION NUMBER: US/09/513,999
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7640
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: UNSURE
; FEATURE:
; LOCATION: 15
; OTHER INFORMATION: Xaa=Ala or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 33
; OTHER INFORMATION: Xaa=Lys or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 73
; OTHER INFORMATION: Xaa=Ala or Asp or Glu or Gly or Val
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 74
; OTHER INFORMATION: Xaa=Ala or Pro or Ser
US-10-793-479-7640

Query Match 36.8%; Score 47.5; DB 6; Length 86;
Best Local Similarity 52.4%; Pred. No. 7.5;
Matches 11; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 1 QRLKYAOEELSNEVLPPPRKM 21
||: ||| ||: ||| |||
Db 44 QRMKRKEELS-KLLPPRRSM 63

RESULT 5
US-10-170-205E-21306
; Sequence 21306, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21306
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-21306

Query Match 36.8%; Score 47.5; DB 6; Length 416;
Best Local Similarity 52.4%; Pred. No. 47;
Matches 11; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 1 QRLKYAOEELSNEVLPPPRKM 21
||: ||| ||: ||| |||
Db 257 QRMKRKEELS-KLLPPRRSM 276

RESULT 6
US-10-170-205E-36793
; Sequence 36793, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36793
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-36793

Query Match 36.8%; Score 47.5; DB 6; Length 538;
Best Local Similarity 52.4%; Pred. No. 63;
Matches 11; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 1 QRLKYAOEELSNEVLPPPRKM 21
||: ||| ||: ||| |||
Db 379 QRMKRKEELS-KLLPPRRSM 398

RESULT 7
US-10-767-701-40723
; Sequence 40723, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

US-60-579-902-9506

APPLICANT: WU, WEI
TITLE OF INVENTION: Nucleotide and Amino Acid Sequences from *Xenorhabdus*
TITLE OF INVENTION: and Uses thereof

FILE REFERENCE: 38-21 (53444) A
CURRENT APPLICATION NUMBER: US/60/579,902
CURRENT FILING DATE: 2004-06-15
NUMBER OF SEQ ID NOS: 14985
SEQ ID NO 10359
LENGTH: 757
TYPE: PRT
ORGANISM: Xenorhabdus bovienii
US-60-579-902-10359

Query Match 35.7%; Score 46; DB 7; Length 757;
Best Local Similarity 45.5%; Pred. No. 1.6e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMK 22
DB 340 ERLVQNSKLVLNKLDFSRKME 361

RESULT 13
US-10-767-701-40468
Sequence 40468, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 40468
LENGTH: 155
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C645_1.pep
US-10-767-701-40468

Query Match 34.9%; Score 45; DB 6; Length 155;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 KYAQEELSNEVLPPPRKMK 21
DB 62 KHAQRIGSPRLPPPNKV 79

RESULT 14
US-10-767-701-56206
Sequence 56206, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 56206
LENGTH: 173
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: 30938965.pep
US-10-767-701-56206

Query Match 34.9%; Score 45; DB 6; Length 173;
Best Local Similarity 36.4%; Pred. No. 40;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 3 LKYAQEELSNEVLPPPRKMKGL 24
DB 54 LNFGADPENNPPLPPPRGRGL 75

RESULT 15
US-10-170-205E-19914
Sequence 19914, Application US/10170205E
GENERAL INFORMATION:
APPLICANT: ADAMS, Mark
TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
FILE REFERENCE: CL001381
CURRENT APPLICATION NUMBER: US/10/170,205E
CURRENT FILING DATE: 2002-06-13
NUMBER OF SEQ ID NOS: 40312
SOFTWARE: PatentIn version 3.2
SEQ ID NO 19914
LENGTH: 185
TYPE: PRT
ORGANISM: Homo sapiens
US-10-170-205E-19914

Query Match 34.9%; Score 45; DB 6; Length 185;
Best Local Similarity 40.0%; Pred. No. 43;
Matches 10; Conservative 6; Mismatches 7; Indels 2; Gaps 1;

QY 1 QRLKYAQEELSNEVLPPP--RKMKG 23
DB 123 EKEKLSKETSEELLPPPVQTQIKG 147

Search completed: July 15, 2004, 15:27:33
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:23:28 ; Search time 16 Seconds
(without alignments)
150.299 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	341	2 S04137	genome polyprotein
2	129	100.0	852	1 GNNYHA	genome polyprotein
3	129	100.0	1358	2 A03905	genome polyprotein
4	129	100.0	2227	1 GNNYHM	genome polyprotein
5	129	100.0	2227	1 GNNYHR	genome polyprotein
6	124	96.1	2227	1 GNNYMK	genome polyprotein
7	123	95.3	2227	1 GNNYHB	genome polyprotein
8	119	92.2	2230	1 GNNYSA	genome polyprotein
9	117	90.7	839	1 GNNYS2	genome polyprotein
10	56.5	43.8	346	2 S74448	regulatory protein
11	52	40.3	859	2 T43701	DNA-directed RNA p
12	52	40.3	1193	2 E88445	protein C26E6.4 [i
13	51	39.5	1119	2 T50995	related to cytoske
14	51	39.5	6642	2 T29757	protein UNC-89 - C
15	50.5	39.1	443	2 E82046	proteinase HelVU,
16	49.5	38.4	159	2 C72210	conserved hypothet
17	49.5	38.4	554	2 A56730	carl protein - Pod
18	49	38.0	55	2 PQ0433	genome polyprotein
19	49	38.0	56	2 PQ0434	genome polyprotein
20	49	38.0	56	2 PQ0428	genome polyprotein
21	49	38.0	56	2 PQ0427	genome polyprotein
22	49	38.0	56	2 PQ0429	genome polyprotein
23	49	38.0	56	2 PQ0432	genome polyprotein
24	49	38.0	56	2 PQ0430	genome polyprotein
25	49	38.0	442	2 A71969	probable histidine
26	49	38.0	465	2 H86482	protein FSJ5.11 [i
27	49	38.0	1174	2 S28976	DNA-directed RNA p
28	48.5	37.6	2278	1 S56274	FAB1 protein - Yea
29	48	37.2	523	2 D85538	hypothetical prote

30 48 37.2 523 2 H30687
31 48 37.2 1176 2 A27826
32 48 37.2 1191 2 S65068
33 48 37.2 1210 2 S35548
34 47.5 36.8 285 2 T15133
35 47.5 36.8 1626 2 T09271
36 47 36.4 223 2 PQ0514
37 47 36.4 254 2 F75575
38 47 36.4 578 2 S03299
39 47 36.4 592 1 L1BY
40 46.5 36.0 361 2 G82530
41 46.5 36.0 486 2 T10089
42 46.5 36.0 634 2 E86293
43 46 35.7 163 2 PC4186
44 46 35.7 262 2 A70428
45 46 35.7 338 2 I56893

ALIGNMENTS

RESULT 1

S04137

genome polyprotein - human hepatitis A virus (strain LCD-1) (fragment)

C:Species: human hepatitis A virus

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000

C:Accession: S04137

R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.

A:Title: Nucleotide sequence of the VPI gene from a Chinese strain of hepatitis A virus

A:Reference number: S04137; MUID:89263805; PMID:2542903

A:Accession: S04137

A:Molecule type: mRNA

A:Residues: 1-341 <AND>

A:Cross-references: EMBL:X14666; NID:G62301; PIDN:CAA32794.1; PID:G4377576

C:Genetics:

A:Gene: VPI

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; polyprotein

F:2-340/Product: coat protein 1D (VPI) #status predicted <MAT>

Query Match 100.0% Score 129; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 9.5e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Db 314 QRLKYAQEELSNEVLPPPRKMKGLF 338

RESULT 2

GNNYHA

genome polyprotein - human hepatitis A virus (strain CR326) (fragment)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03904

R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W

A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.

A:Reference number: A03904; MUID:85185648; PMID:2985793

A:Accession: A03904

A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:G329592; PIDN:AAA45470.1; PID:G329593

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 96.1%; Score 124; DB 1; Length 2227;
Best Local Similarity 96.0%; Pred. No. 4.6e-10;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
|||||:|||||:|||||:|||||:|||||
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 7
GNNYHB
Genome polyprotein - human hepatitis A virus (strain MBB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhardt, Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolated from a patient with acute hepatitis A) [cDNA] (GenBank)
A:Reference number: J50303; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrophobic; RNA-directed RNA polymerase
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 95.3%; Score 123; DB 1; Length 2227;
Best Local Similarity 96.0%; Pred. No. 6.6e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
|||||:|||||:|||||:|||||:|||||
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 8
GNNYSA
Genome polyprotein - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S03965
R:Tzarev, S.A.
submitted to JPIPD, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:g222597; PIDN:BAA00766.1; PID:g222598
R:Tzarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and sequence with other HAV strains
A:Reference number: J01080; PMID:91311420; PMID:1649901
A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhikov
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL1>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhikov
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian hepatitis A virus
A:Reference number: S03965; PMID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; genome polyprotein
F:1-27/Product: coat protein 1A #status predicted <C1A>
F:28-249/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:496-795/Product: coat protein 1D #status predicted <C1D>
F:796-984/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: protein 3A #status predicted <P3A>
F:1499-1521/Product: protein 3B #status predicted <P3B>
F:1522-1741/Product: protein 3C #status predicted <P3C>
F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 92.2%; Score 119; DB 1; Length 2230;
Best Local Similarity 88.0%; Pred. No. 2.7e-09;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
|||||:|||||:|||||:|||||:|||||
DB 814 QRLKYAQEELSNEVLPPPRKMKGLF 838

RESULT 9
GNNYS2
Genome polyprotein - simian hepatitis A virus (strain CV-145) (fragment)
N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat protein 1C; species: simian hepatitis A virus
C:Species: simian hepatitis A virus
A:Note: host Macaca fascicularis (cynomolgus macaque)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999
C:Accession: JQ1180
R:Nainan, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.
J. Gen. Virol. 72, 1685-1689, 1991
A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus macaque
A:Reference number: JQ1180; PMID:91311421; PMID:1649902
A:Accession: JQ1180
A:Molecule type: genomic RNA
A:Residues: 1-839 <NAI>
A:Cross-references: GB:M59286; NID:g329599; PIDN:AAA45473.1; PID:g555083
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-23/Product: coat protein 1A #status predicted <VP0>
F:24-245/Product: coat protein 1B #status predicted <VP3>
F:246-491/Product: coat protein 1C #status predicted <VP1>
F:492-839/Product: core protein 2A (fragment) #status predicted <P2P>
F:261,312,728,756/Binding site: carbohydrate (Asn) #status predicted

Query Match 90.7%; Score 117; DB 1; Length 839;
Best Local Similarity 84.0%; Pred. No. 1.8e-09;
Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
|||||:|||||:|||||:|||||:|||||
DB 809 QRLKYAQEELSNEVLPPPRKMKGLF 833

RESULT 10

Qy 1 QRLKYAQBELSNEVLPPPRKMK 22
:|: :||| | | : || |
Db 1353 RRVSFARBEELPKVEIDSDRKKK 1374

RESULT 15
E82046
proteinase HslVU, ATPase subunit HslU VC2674 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: E82046
R:Heidelbergl, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: E82046
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <HEI>
A:Cross-references: GB:AE004333; GB:AE003852; NID:g9657266; PIDN:AAF95815.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2674
A:Map position: 1
C:Superfamily: heat shock protein hslU; FtsH/SEC18/CDC48-type ATP-binding domain homolog

Query Match 39.1%; Score 50.5; DB 2; Length 443;
Best Local Similarity 39.3%; Pred. No. 10;
Matches 11; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

Qy 1 QRLKYAQBELSNE-----VLPPPRKMK 23
:|: :||| | | : || |
Db 117 EKVFRABELAEERVLDAIPLPPRDAWG 144

Search completed: July 15, 2004, 15:27:03
Job time : 17 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:18:18 ; Search time 14 Seconds
(without alignments)
92.982 Million cell updates/sec

Title: US-09-171-432A-47
Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	341	1	POLG_HPAV1
2	129	100.0	852	1	P06442 hepatitis a
3	129	100.0	2226	1	POLG_HPAV2
4	129	100.0	2226	1	POLG_HPAV4
5	129	100.0	2226	1	POLG_HPAV8
6	129	100.0	2227	1	POLG_HPAVH
7	129	100.0	2227	1	POLG_HPAVL
8	123	95.3	2227	1	POLG_HPAVM
9	119	92.2	2230	1	POLG_HPAVS
10	117	90.7	839	1	POLG_HPAVT
11	103	79.8	808	1	POLG_HPAVG
12	53	41.1	636	1	NAF1_HUMAN
13	52	40.3	1193	1	RPB2_CABEL
14	51	39.5	6632	1	UNB9_CABEL
15	50.5	39.1	443	1	HSLU_VIBCH
16	50.5	39.1	443	1	HSLU_VIBVU
17	49.5	38.4	554	1	PEX2_PODAN
18	49	38.0	1174	1	RPB2_HUMAN
19	48.5	37.6	2278	1	FAB1_YEAST
20	48	37.2	1176	1	RPB2_DROME
21	48	37.2	1191	1	RPB2_LYCES
22	48	37.2	1210	1	RPB2_SCHPO
23	47	36.4	578	1	HEMA_INBME
24	47	36.4	592	1	ABP1_YEAST
25	46.5	36.0	399	1	BMA8_MOUSE
26	46	35.7	338	1	TAP4_HUMAN
27	46	35.7	480	1	11SB_CUCMA
28	46	35.7	574	1	HEMA_INBMD
29	46	35.7	576	1	HEMA_INBUS
30	46	35.7	578	1	HEMA_INBVI
31	46	35.7	583	1	HEMA_INBEN
32	46	35.7	583	1	HEMA_INBOR
33	46	35.7	583	1	HEMA_INBSI

34 46 35.7 585 1 HEMA_INBRE P17504 influenza b
35 46 35.7 585 1 HEMA_INBYK P23092 influenza b
36 46 35.7 1085 1 CARB_HELFU Q98kt2 helicobac
37 45 34.9 315 1 RSEB_HAEIN P44792 haemophilu
38 45 34.9 445 1 ML64_HUMAN Q14849 homo sapien
39 45 34.9 733 1 HEXA_BLADI Q17127 blaberus di
40 45 34.9 1188 1 RPB2_ARATH P38420 arabidopsis
41 45 34.9 1312 1 LI59_CABEL O44757 caenorhadi
42 44.5 34.5 702 1 CRN_DROME P17886 drosophila
43 44.5 34.5 3678 1 DMD_MOUSE P11531 mus musculu
44 44.5 34.5 3680 1 DMD_CANFA O97592 canis famil
45 44.5 34.5 3685 1 DMD_HUMAN P11532 homo sapien

ALIGNMENTS

RESULT 1
POLG_HPAV1
ID POLG_HPAV1 STANDARD; PRT; 341 AA.
AC P13672;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP3 (1C); Coat protein VP1
DE (1D); Core protein P2A] (Fragment).
OS Hepatitis A virus (strain LCD-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
RT hepatitis A virus (HAV).";
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14666; CAA32794.1; -.
CC PIR; S04137; S04137.
CC InterPro; IPR008975; Viral_cap_coat.
KW Polyprotein; Coat protein; Core protein.
FT NON_TER 1 1
FT CHAIN <1 1 COAT PROTEIN VP3.
FT CHAIN 2 340 COAT PROTEIN VP1.
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;
Query Match 100.0%; Score 129; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 6.7e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;
Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 314 QRLKYAQEELSNEVLPPPRKMKGLF 338
RESULT 2
POLG_HPAV1
ID POLG_HPAV1 STANDARD; PRT; 852 AA.

RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA](N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59809; AAA45469.1; -;
 DR MEROPS; C03.005; -;
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 DR Polyprotein; Coat protein; Core protein; Transferase;
 DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 KW CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;
 Query Match 100.0%; Score 129; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 5.2e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834
 RESULT 5
 POLG HPVAV8 STANDARD; PRT; 2226 AA.
 AC P26582;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain 18f).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 RN NCBI_TaxID=12096;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Croweans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA](N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59808; AAA45467.1; -;
 DR PDB; 1QA7; 15-MAY-00.
 DR MEROPS; C03.005; -;
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 DR Polyprotein; Coat protein; Core protein; Transferase;
 DR RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 KW CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;
 Query Match 100.0%; Score 129; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 5.2e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
 DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834
 RESULT 6
 POLG HPVAVH STANDARD; PRT; 2227 AA.
 AC P08617; P06443; Q81082;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain HM-175).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12098;

RN SEQUENCE FROM N.A.
RP STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus;
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses";
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Attenuated;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus;
RT comparison with wild-type virus";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166283; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14114; AAA45475.1; --
CC EMBL; M14707; AAA45465.1; --
CC EMBL; M14707; AAA45466.1; ALT_INIT.
CC EMBL; M16632; AAA45471.1; --
CC PIR; A03905; A03905.
CC PIR; A25981; GNNYHM.
CC PIR; A94149; GNNYMK.
CC FDB; LHAV; 23-DEC-96.
CC MEROPS; C03.005; --
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.
KW CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
FT CHAIN 492 836 COAT PROTEIN VP3 (PIC).
FT CHAIN 837 COAT PROTEIN VP1 (PID).
FT CHAIN 837 CORE PROTEIN P2A.

FT CHAIN 981 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
FT VARIANT 77 77 K -> R (IN ATTENUATED STRAIN).
FT VARIANT 764 764 E -> V (IN ATTENUATED STRAIN).
FT VARIANT 821 821 N -> S (IN ATTENUATED STRAIN).
FT VARIANT 1052 1052 A -> V (IN ATTENUATED STRAIN).
FT VARIANT 1062 1062 G -> A (IN ATTENUATED STRAIN).
FT VARIANT 1118 1118 K -> M (IN ATTENUATED STRAIN).
FT VARIANT 1151 1151 E -> K (IN ATTENUATED STRAIN).
FT VARIANT 1163 1163 F -> S (IN ATTENUATED STRAIN).
FT VARIANT 1277 1277 H -> I (IN ATTENUATED STRAIN).
FT VARIANT 1500 1500 V -> Y (IN ATTENUATED STRAIN).
FT VARIANT 1805 1805 D -> N (IN ATTENUATED STRAIN).
FT VARIANT 1930 1930 S -> T (IN ATTENUATED STRAIN).
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;
Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 810 QRLKYAQEELSNEVLPPPRKMKGLF 834
RESULT 7
POLG HPVAVL STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain LA).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12099;
RX SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127; Potter S.J., Renard A.,
RA Najarian R., Caput D., Gee W.W., Merryweather J., van Nest G., Dina D.;
RA "Primary structure and gene organization of human hepatitis A virus";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K02990; AAA45472.1; --
CC PIR; A03903; GNNYHR.
CC MEROPS; C03.005; --
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC Polyprotein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.
KW CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
FT CHAIN 492 836 COAT PROTEIN VP3 (PIC).
FT CHAIN 837 COAT PROTEIN VP1 (PID).
FT CHAIN 837 CORE PROTEIN P2A.

```

DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
KW Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 8
POLG HPVAVM STANDARD; PRT; 2227 AA.
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain MB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wessel T., Kiehn R., Wimmer E.,
RA Deinhardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
RT virus (isolate MB).";
RL Virus Res. 8:153-171(1987).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20273; AAA45474.1; -
DR MEROPS; C03.005; -

```

```

DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
KW Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 95.3%; Score 123; DB 1; Length 2227;
Best Local Similarity 96.0%; Pred. No. 4.2e-10;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 810 QRLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 9
POLG HPVAVM STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D00924; BAA00766.1; -
 DR EMBL; X15461; CAA33490.1; -
 DR PIR; A30470; GNNYS2.
 DR MEROPS; C03.005; -
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys_Ser_trypsin.
 DR InterPro; IPR000605; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_P5vir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 27 COAT PROTEIN VP4 (PIA).
 FT CHAIN 28 249 COAT PROTEIN VP2 (PIB).
 FT CHAIN 250 495 COAT PROTEIN VP3 (PIC).
 FT CHAIN 496 795 COAT PROTEIN VP1 (PID).
 FT CHAIN 796 984 CORE PROTEIN P2A.
 FT CHAIN 985 1091 CORE PROTEIN P2B.
 FT CHAIN 1092 1426 CORE PROTEIN P2C.
 FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
 FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
 FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
 FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
 FT CHAIN 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;
 SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 92.2%; Score 119; DB 1; Length 2230;
 Best Local Similarity 88.0%; Pred. No. 1.7e-09;
 Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAEELSNEVLPPPRKMGFLF 25
 DB 814 QRLKYAEELSNEVLPPPRKMGFLF 838

RESULT 10
 POLG HPVAV STANDARD; PRT; 839 AA.
 AC P31788;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
 OS Simian hepatitis A virus (strain CV-145).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=31707;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91311421; PubMed=1649902;
 RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
 RT "Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus macaques (Macaca fascicularis).";
 RL J. Gen. Virol. 72:1685-1689(1991).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M59286; AAA45473.1; -
 DR PIR; J01180; GNNYS2.
 DR InterPro; IPR008975; Viral_cap_coat.
 KW Polyprotein; Coat protein; Core protein.
 FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
 FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
 FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
 FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
 FT CHAIN ? >839 CORE PROTEIN P2A.
 FT NON_TER 839 839
 SQ SEQUENCE 839 AA; 93925 MW; 2CACC4BD1E192DEC CRC64;

Query Match 90.7%; Score 117; DB 1; Length 839;
 Best Local Similarity 84.0%; Pred. No. 1.2e-09;
 Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAEELSNEVLPPPRKMGFLF 25
 DB 809 QRLKYAEELSNEVLPPPRKMGFLF 833

RESULT 11
 POLG HPVAV STANDARD; PRT; 808 AA.
 AC Q02381;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
 OS Hepatitis A virus (strain GA76).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=31706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92260183; PubMed=1316423;
 RA Khanna B., Speilbring J.E., Innis B.L., Robertson B.H.;
 RT "Characterization of a genetic variant of human hepatitis A virus.";
 RL J. Med. Virol. 36:118-124(1992).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M66695; AAA45477.1; -
 DR InterPro; IPR008975; Viral_cap_coat.
 KW Polyprotein; Coat protein; Core protein.
 FT NON_TER 1 1
 FT CHAIN <1 2 COAT PROTEIN VP4 (PIA).
 FT CHAIN 3 223 COAT PROTEIN VP2 (PIB).
 FT CHAIN 224 470 COAT PROTEIN VP3 (PIC).
 FT CHAIN 471 770 COAT PROTEIN VP1 (PID).
 FT CHAIN 771 >808 CORE PROTEIN P2A.
 FT NON_TER 808 808
 SQ SEQUENCE 808 AA; 90632 MW; D80CE7B57A479C12 CRC64;

Query Match 79.8%; Score 103; DB 1; Length 808;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 QRLKYAQAEELSNEVLPPPRK 20
Db |||||
789 QRLKYAQAEELSNEVLPPPRK 808

RESULT 12
ID NAF1 HUMAN STANDARD; PRT; 636 AA.
AC Q15025; O76008; Q96L9; Q99833; Q9HJ3;
AT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nef-associated factor 1 (Naf1) (HIV-1 Nef interacting protein)
DE (Varion-associated nuclear shuttling protein) (VAN) (hVAN) (TNFAIP3
DE interacting protein 1).
GN TNIP1 OR NAF1 OR KIAA0113.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Peripheral blood;
RX MEDLINE=99120485; PubMed=9923610;
RA Fukushima M., Dixon J., Kimura T., Tsurutani N., Dixon M.J.,
RA Yamamoto N.;
RT Identification and cloning of a novel cellular protein Naf1, Nef-
RT associated factor 1, that increases cell surface CD4 expression." ;
RN PERS Lett. 442:83-88(1999).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Leukocyte;
RX MEDLINE=20541981; PubMed=11090181;
RA Gupta K., Ott D., Hope T.J., Silliciano R.F., Boeke J.D.;
RT "A human nuclear shuttling protein that interacts with human
RT immunodeficiency virus type 1 matrix is packaged into virions." ;
RN J. Virol. 74:11811-11824(2000).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muhlath S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences." ;
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE OF 136-636 FROM N.A. (ISOFORM 2).
RC TISSUE=Craniofacial;
RX MEDLINE=96276047; PubMed=8681136;
RA Loftus S.K., Dixon J., Koprivnikar K., Dixon M.J., Wasmuth J.J.;
RT "Transcriptional map of the Treacher Collins candidate gene region." ;
RN Genome Res. 6:26-34(1996).
RN [5]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RA Ohara O., Nagase T., Kikuno R., Nomura N.;
RN Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

```

[6]
SEQUENCE OF 341-636 FROM N.A. (ISOFORM 1).
TISSUE=Bone marrow;
MEDLINE=95308325; PubMed=778527;
Nagase T., Miyajima N., Tanaka A., Sazuka T., Seki N., Sato S.,
Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;
"Prediction of the coding sequences of unidentified human genes. III.
The coding sequences of 40 new genes (KIAA0081-KIAA0120) deduced by
analysis of cDNA clones from human cell line KG-1." ;
DNA Res. 2:37-43(1995).

[7]
SEQUENCE OF 94-412 FROM N.A.
Fukushi M., Kimura T., Yamamoto N.;
Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: Interacts with zinc finger protein A20/TNFAIP3 and
inhibits TNF-induced NF-kappa-B-dependent gene expression by
interfering with an RIP- or TRAF2-mediated transactivation signal
(By similarity). Increases cell surface CD4(T4) antigen
expression. Interacts with HIV-1 matrix protein and is packaged
into virions and overexpression can inhibit viral replication. May
regulate matrix nuclear localization, both nuclear import of PIC
(preintegration complex) and export of GAG polyprotein and viral
genomic RNA during virion production.
-!- SUBUNIT: Interacts with TNFAIP3 (By similarity). Interacts with
HIV-1 matrix protein.
-!- SUBCELLULAR LOCATION: Cytoplasmic. Shuttles between the nucleus
and cytoplasm in a CRM1-dependent manner.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1; Synonyms=Alpha;
IsoId=Q15025-1; Sequence=Displayed;
Name=2; Synonyms=Beta;
IsoId=Q15025-2; Sequence=VSP_003913;
Note=No experimental confirmation available;
-!- TISSUE SPECIFICITY: Ubiquitous. Strongly expressed in peripheral
blood lymphocytes, spleen and skeletal muscle, and is weakly
expressed in the brain.
-!- CAUTION: Ref.7 sequence differs from that shown due to frameshifts
in positions 152 and 154.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.

EMBL; AJ011895; CAA09855.1; -
EMBL; AJ011896; CAA09856.1; -
EMBL; AY012155; AAG42154.1; -
EMBL; BC012133; AAH12133.1; -
EMBL; BC014008; AAH14008.1; -
EMBL; U39403; AAC99999.1; -
EMBL; D30755; BAA06416.2; -
EMBL; U83844; AAB41438.1; ALT_FRAME.
EMBL; HGNC:16903; TNIP1.
MIM; 607714; -
GO; GO:0005622; C:intracellular; TAS.
GO; GO:0005515; F:protein binding; TAS.
GO; GO:0009101; P:glycoprotein biosynthesis; IDA.
GO; GO:0045071; P:negative regulation of viral genome replica. . .; TAS.
GO; GO:0045071; P:negative regulation of viral genome replica. . .; TAS.
KW Coiled coil; Nuclear protein; Alternative splicing.
FT DOMAIN 20 73
FT COILED COIL (POTENTIAL).
FT DOMAIN 196 258
FT COILED COIL (POTENTIAL).
FT DOMAIN 294 535
FT COILED COIL (POTENTIAL).
FT DOMAIN 94 412
FT INTERACTS WITH NEF.
FT DOMAIN 524 530
FT NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 539 636
FT PRO-RICH.
FT VARSPLIC 627 636
FT SPKNDREGFPQ -> PADLRLPN (in isoform 2).
FT /FTID=VSP_003913.
FT CONFLICT 148 148
FT CONFLICT 299 299
FT G -> D (IN REF. 3; AAH12133).
FT A -> P (IN REF. 2).

```
SQ SEQUENCE 636 AA; 71864 MW; D81B96BEAD50D871 CRC64;
Query Match 41.1%; Score 53; DB 1; Length 636;
Best Local Similarity 50.0%; Pred No. 3.7;
Matches 13; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY 2 RLKQAQEEEL--SNEVLPPPRKMGFLF 25
   ||| ||| ||| ||| |||
DB 58 RLKQAQEEELVKNELLPPPSPLGSLF 83

RESULT 13
RP02 CAEEL STANDARD; PRT; 1193 AA.
ID RP02 CAEEL STANDARD; PRT; 1193 AA.
AC Q10578;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE DNA-directed RNA polymerase II second largest subunit (EC 2.7.7.6)
DE (RNA polymerase I subunit 2).
DE C26E6..4.
GN Caenorhabditis elegans.
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Fulton L.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE OF 200-1058 FROM N.A.
RX MEDLINE=95041334; PubMed=7953533;
RA Sidow A., Thomas W.K.;
RL "A molecular evolutionary framework for eukaryotic model organisms.";
Curr. Biol. 4:596-603(1994).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
of DNA into RNA using the four ribonucleoside triphosphates as
substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}(N).
CC -!- SUBUNIT: Each class of RNA polymerase is assembled from 9 to 14
different polypeptides. This subunit is the second largest
component of RNA polymerase II.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
found in eukaryotic nuclei: polymerase I for the ribosomal RNA
precursor, polymerase II for the mRNA precursor, and polymerase
III for 5S and tRNA genes.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta chain family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenses@isb-sib.ch).
-----
EMBL; U13875; AAA021158.1; --
EMBL; U10333; AAA50224.1; --
PIR; E88445; E88445.
PIR; T43701; T43701.
WormPep; C26E6.4; CE01162.
InterPro; IPR007121; RNA_pol_B.
InterPro; IPR007644; RNA_pol_Rpb2_1.
InterPro; IPR007642; RNA_pol_Rpb2_2.
InterPro; IPR007645; RNA_pol_Rpb2_3.
InterPro; IPR007646; RNA_pol_Rpb2_4.
InterPro; IPR007647; RNA_pol_Rpb2_5.
InterPro; IPR007120; RNA_pol_Rpb2_6.
InterPro; IPR007641; RNA_pol_Rpb2_7.
Pfam; PF04563; RNA_pol_Rpb2_1; 1.

DR Pfam; PF04561; RNA_pol_Rpb2_2; 1.
DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
DR Pfam; PF00562; RNA_pol_Rpb2_6; 1.
DR Pfam; PF04560; RNA_pol_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein.
FT ZN_FING 1125 1146 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1193 AA; 134904 MW; B8A85E74E9CC7EBE CRC64;

Query Match 40.3%; Score 52; DB 1; Length 1193;
Best Local Similarity 62.5%; Pred No. 10;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 QRLKQAQEEELSNEVLFP 16
   ||| ||| ||| ||| |||
DB 340 QRIKYAREILQKELLP 355

RESULT 14
UN89 CAEEL STANDARD; PRT; 6632 AA.
ID UN89 CAEEL STANDARD; PRT; 6632 AA.
AC Q01761; Q17362;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).
DE UNC-89 OR C09D1.1.
GN Caenorhabditis elegans.
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.
RP STRAIN=Bristol N2;
RX MEDLINE=96180278; PubMed=8603916;
RA Benian G.M., Tinley T.L., Tang X., Borodovsky M.;
RL "The Caenorhabditis elegans gene unc-89, required for muscle M-line
assembly, encodes a giant modular protein composed of Ig and signal
transduction domains.";
J. Cell Biol. 132:835-848(1996).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Du Z., Le T.T., Wilson R.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
[3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Structural component of the muscle M-line. Myofilament
lattice assembly begins with positional cues laid down in the
basement membrane and muscle cell membrane. UNC-89 responds to
these signals, localizes, and then participates in assembling an
M-line.
CC -!- TISSUE SPECIFICITY: Localizes to the middle of A-bands.
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC -!- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 5 RCD5 domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to licenses@isb-sib.ch).
-----
```


FT	DOMAIN	5067	5160	IG-LIKE C2-TYPE 37.
FT	DOMAIN	5171	5260	IG-LIKE C2-TYPE 38.
FT	DOMAIN	5277	5366	IG-LIKE C2-TYPE 39.
FT	DOMAIN	5383	5472	IG-LIKE C2-TYPE 40.
FT	DOMAIN	5487	5578	IG-LIKE C2-TYPE 41.
FT	DOMAIN	5595	5685	IG-LIKE C2-TYPE 42.
FT	DOMAIN	5701	5790	IG-LIKE C2-TYPE 43.
FT	DOMAIN	5815	5904	IG-LIKE C2-TYPE 44.
FT	DOMAIN	6038	6130	IG-LIKE C2-TYPE 45.
FT	DOMAIN	6150	6239	IG-LIKE C2-TYPE 46.
FT	DOMAIN	6275	6368	FIBONECTIN TYPE-III.
FT	DOMAIN	6413	6502	IG-LIKE C2-TYPE 48.
FT	DOMAIN	6507	6596	IG-LIKE C2-TYPE 49.
FT	DISULFID	568	621	POTENTIAL.
FT	DISULFID	2908	2975	POTENTIAL.
FT	DISULFID	3015	3065	POTENTIAL.
FT	DISULFID	3707	3759	POTENTIAL.
FT	DISULFID	3826	3890	POTENTIAL.
FT	DISULFID	5092	5157	POTENTIAL.
FT	DISULFID	5298	5350	POTENTIAL.
FT	DISULFID	5508	5560	POTENTIAL.
FT	DISULFID	5616	5669	POTENTIAL.
FT	DISULFID	5722	5764	POTENTIAL.
FT	DISULFID	5836	5901	POTENTIAL.
FT	DISULFID	5946	5998	POTENTIAL.
FT	DISULFID	6036	6171	POTENTIAL.
FT	DISULFID	6421	6486	POTENTIAL.
FT	CONFLICT	2137	2137	A -> P (IN REF. 1).
FT	CONFLICT	2245	2247	AKA -> PKP (IN REF. 1).
FT	CONFLICT	2258	2258	A -> P (IN REF. 1).
FT	CONFLICT	2284	2284	E -> G (IN REF. 1).
FT	CONFLICT	2297	2297	M -> I (IN REF. 1).
FT	CONFLICT	3531	3531	A -> G (IN REF. 1).
FT	CONFLICT	3884	3888	DAGEY -> RRRRI (IN REF. 1).
FT	CONFLICT	3929	3929	A -> V (IN REF. 1).
FT	CONFLICT	5134	5134	A -> P (IN REF. 1).
FT	CONFLICT	5145	5145	T -> S (IN REF. 1).
FT	CONFLICT	5185	5185	G -> A (IN REF. 1).
FT	CONFLICT	5199	5199	K -> N (IN REF. 1).
FT	CONFLICT	5202	5202	L -> F (IN REF. 1).
FT	CONFLICT	5213	5213	F -> L (IN REF. 1).
FT	CONFLICT	6178	6178	A -> G (IN REF. 1).
FT	CONFLICT	6268	6268	K -> E (IN REF. 1).
SQ	SEQUENCE	6632 AA;	731665 MW;	262D3EDD62960B89 CRC64;
 Query Match 39.5%; Score 51; DB 1; Length 6632; Best Local Similarity 45.5%; Pred.No. 96; Matches 10; Conservative 5; Mismatches 7; Indels				
Qy	1 QRLKYAOEELSNEVLPPPRMKK 22 :: :: :			
Db	1353 RVSFAEELPKVIDSDRKKK 1374			
 RESULT 15				
HSLU	VIBCH			
ID	_HSLU_VIBCH	STANDARD;	PRT;	443 AA.
AC	Q9NQN7;			
DT	16-OCT-2001	(Rel. 40,	Created)	
DT	16-OCT-2001	(Rel. 40,	Last sequence update)	
DE	10-OCT-2003	(Rel. 42,	Last annotation update)	
DE	ATP-dependent hsl protease ATP-binding subunit hslU.			
GN	HSLU OR VC2674.			
OS	Vibrio cholerae.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;			
OC	Vibrionaceae; Vibrio.			
OX	NCBI_TaxID=666;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=E1 Tor N16961 / Serotype O1;			
RA	MEDLINE=20406833; PubMed=10952301;			
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwi-			

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:23:03 ; Search time 36 Seconds
(without alignments)
219.110 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	94	12	Q9ENT9
2	129	100.0	94	12	Q9ENU3
3	129	100.0	94	12	Q9ENV9
4	129	100.0	94	12	Q9ENV5
5	129	100.0	94	12	Q9ENV2
6	129	100.0	94	12	Q9ENV1
7	129	100.0	94	12	Q9ENT1
8	129	100.0	94	12	Q9ENV8
9	129	100.0	94	12	Q9ENV2
10	129	100.0	94	12	Q9ENV5
11	129	100.0	94	12	Q9ENV1
12	129	100.0	94	12	Q9ENV9
13	129	100.0	94	12	Q9ENV7
14	129	100.0	94	12	Q9ENV8
15	129	100.0	94	12	Q9ENV6
16	129	100.0	94	12	Q9ENV0

17	129	100.0	94	12	Q9ENU6	Q9enu6 hepatitis a
18	129	100.0	94	12	Q9ENU4	Q9enu4 hepatitis a
19	129	100.0	94	12	Q9ENW1	Q9enw1 hepatitis a
20	129	100.0	94	12	Q9ENT5	Q9ent5 hepatitis a
21	129	100.0	94	12	Q9ENV7	Q9enu7 hepatitis a
22	129	100.0	94	12	Q9ENV4	Q9env4 hepatitis a
23	129	100.0	115	12	Q9DWR4	Q9dwr4 hepatitis a
24	129	100.0	116	12	Q92941	Q92941 hepatitis a
25	129	100.0	116	12	Q9W7X7	Q9w7x7 hepatitis a
26	129	100.0	116	12	Q9W7S7	Q9w7s7 hepatitis a
27	129	100.0	116	12	Q71977	Q71977 hepatitis a
28	129	100.0	116	12	Q8B8K6	Q8b8k6 hepatitis a
29	129	100.0	116	12	Q8B8K5	Q8b8k5 hepatitis a
30	129	100.0	116	12	Q8B8K4	Q8b8k4 hepatitis a
31	129	100.0	116	12	Q8B8K3	Q8b8k3 hepatitis a
32	129	100.0	126	12	Q98W19	Q98w19 hepatitis a
33	129	100.0	132	12	Q8V4L9	Q8v4l9 hepatitis a
34	129	100.0	132	12	Q8V4L6	Q8v4l6 hepatitis a
35	129	100.0	132	12	Q8V4J9	Q8v4j9 hepatitis a
36	129	100.0	132	12	Q8V4J5	Q8v4j5 hepatitis a
37	129	100.0	132	12	Q8V4M5	Q8v4m5 hepatitis a
38	129	100.0	132	12	Q8V4M7	Q8v4m7 hepatitis a
39	129	100.0	132	12	Q8V4N2	Q8v4n2 hepatitis a
40	129	100.0	132	12	Q8V4M4	Q8v4m4 hepatitis a
41	129	100.0	132	12	Q8V4J7	Q8v4j7 hepatitis a
42	129	100.0	132	12	Q8V4L5	Q8v4l5 hepatitis a
43	129	100.0	132	12	Q8V4K3	Q8v4k3 hepatitis a
44	129	100.0	132	12	Q8V4L0	Q8v4l0 hepatitis a
45	129	100.0	132	12	Q8V4M2	Q8v4m2 hepatitis a

ALIGNMENTS

RESULT 1

Q9ENT9	PRELIMINARY;	PRT;	94 AA.
AC	Q9ENT9;		
DT	01-MAR-2001 (Tremblrel. 16, Created)		
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	Polyprotein (Fragment).		
OS	Hepatitis A virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;		
OC	Hepatovirus.		
OX	NCBI_TaxID=12092;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Nagasaki 32;		
RA	Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,		
RA	Yasuoka A., Oka S.;		
RT	"Influence of HIV-1 infection on acute hepatitis A virus (HAV)		
RT	infection: Prolonged HAV viremia and mild liver injury."		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB038298; BAB11836.1; --		
DR	PIR; PQ0427; PQ0427.		
DR	PIR; PQ0428; PQ0428.		
DR	PIR; PQ0430; PQ0430.		
FT	NON TER 1		
FT	NON TER 94		
SQ	SEQUENCE 94 AA; 10807 MW; F194CE91BB8C4AFD CRC64;		

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

Db 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

RESULT 2

Q9ENU3

```

ID Q9ENV3 PRELIMINARY; PRT; 94 AA.
AC Q9ENV3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 27;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038294; BAB11832.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10892 MW; 9AF8BE91BB8C53CC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 3
Q9ENV9 PRELIMINARY; PRT; 94 AA.
AC Q9ENV9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 03;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038278; BAB11816.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10862 MW; 9AF9EFD4AE8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 4

```

```

Q9ENV5 PRELIMINARY; PRT; 94 AA.
AC Q9ENV5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 08;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038282; BAB11820.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 5
Q9ENV2 PRELIMINARY; PRT; 94 AA.
AC Q9ENV2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 28;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038295; BAB11833.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

```

```
RESULT 6
Q9ENT1
ID Q9ENT1 PRELIMINARY; PRT; 94 AA.
AC Q9ENT1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kantou 49;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038306; BAB11844.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10842 MW; 741850915B86E1A6 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

RESULT 7
Q9ENT1
ID Q9ENT1 PRELIMINARY; PRT; 94 AA.
AC Q9ENT1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 29;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038296; BAB11834.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

RESULT 8
Q9ENV2
ID Q9ENV2 PRELIMINARY; PRT; 94 AA.
AC Q9ENV2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 22;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038283; BAB11827.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMKGLF 85

RESULT 9
Q9ENV2
ID Q9ENV2 PRELIMINARY; PRT; 94 AA.
AC Q9ENV2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 15;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038285; BAB11823.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
DB 61 QRLKYAQEELSNEVLPPPRKMKGLF 85
```

Best Local Similarity 100.0%; Pred. No. 3.8e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 25; Conservative 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 10

Q9ENU5 Q9ENU5 PRELIMINARY; PRT; 94 AA.
AC Q9ENU5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 25;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038292; BAB11824.1; -;
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 11

Q9ENV1 Q9ENV1 PRELIMINARY; PRT; 94 AA.
AC Q9ENV1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 18;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038286; BAB11824.1; -;
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94

SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 12

Q9ENU9 Q9ENU9 PRELIMINARY; PRT; 94 AA.
AC Q9ENU9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 21;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038288; BAB11826.1; -;
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGFLF 85

RESULT 13

Q9ENV7 Q9ENV7 PRELIMINARY; PRT; 94 AA.
AC Q9ENV7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 06;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038280; BAB11818.1; -;
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.

```
DR PIR; PQ0431; PQ0431. 1
FT NON_TER 1
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGGLF 85

RESULT 14
Q9ENV8
ID Q9ENV8 PRELIMINARY; PRT; 94 AA.
AC Q9ENV8; 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 04;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038279; BAB11817.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGGLF 85

RESULT 15
Q9ENV6
ID Q9ENV6 PRELIMINARY; PRT; 94 AA.
AC Q9ENV6; 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 07;
RA Ida S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038281; BAB11819.1; -.
DR PIR; PQ0427; PQ0427.
```

```
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431. 1
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRLKYAQEELSNEVLPPPRKMGGLF 25
Db 61 QRLKYAQEELSNEVLPPPRKMGGLF 85

Search completed: July 15, 2004, 15:26:36
Job time : 37 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:36:31 ; Search time 40.1333 Seconds
(without alignments)
140.804 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101
Sequence: 1 VLPFRKMGFLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 581562

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	20	AAW42925	Immunogen
2	101	100.0	21	AAB69442	Synthetic
3	65	64.4	25	AAW42930	Immunogen
4	65	64.4	26	AAB69447	Synthetic
5	45	44.6	23	AAR15517	BMP-8 pep
6	42	41.6	20	AAW42926	Immunogen
7	42	41.6	21	AAB69443	Synthetic
8	39	38.6	20	AAW42924	Immunogen
9	39	38.6	21	AAB69441	Synthetic
10	37	36.6	27	ABB50903	Human sec
11	37	36.6	27	ABO45160	Novel hum
12	37	36.6	27	ABO26640	Protein a
13	36	35.6	28	AAB13951	Antiviral
14	36	35.6	29	AAU08313	Human his
15	35	34.7	28	AAB13952	Antiviral
16	35	34.7	29	AAU08314	Human his
17	34	33.7	15	AAB38287	Human sec
18	34	33.7	15	AAB38291	Gene 7 hu
19	34	33.7	15	AAB39297	Gene 43 h
20	34	33.7	15	AAB38184	Gene 9 hu
21	34	33.7	15	AAB51822	Gene 44 h
22	34	33.7	15	AAB75310	Gene 15 h
23	34	33.7	15	AAG62241	Human sec
24	34	33.7	17	AAB34679	Gene 44 h
25	34	33.7	17	AAB44387	Sequence

26	34	33.7	17	4	AAB64980	Gene 41 h
27	34	33.7	18	3	AAB51896	Gene 17 h
28	34	33.7	18	3	AAB27644	Human sec
29	34	33.7	18	3	AAB36538	Human ser
30	34	33.7	18	3	AAB27787	Sequence
31	34	33.7	18	3	AAB36545	Human ser
32	34	33.7	19	2	AAW65529	Hepatitis
33	33.5	33.2	20	4	AAO09220	Human pol
34	33	32.7	15	5	ABP71855	Human euk
35	33	32.7	16	2	AAW25413	Crk N-ter
36	33	32.7	16	4	AAB67781	Peptide f
37	33	32.7	25	6	ABU07003	Maize Sta
38	33	32.7	30	2	AAB67424	HCV non-s
39	33	32.7	30	2	AAW67423	HCV non-s
40	33	32.7	30	2	AAW67422	HCV non-s
41	33	32.7	30	2	AAW67425	HCV non-s
42	33	32.7	30	4	AAW38506	Peptide #
43	33	32.7	30	4	AAW78242	Human bon
44	33	32.7	30	4	AAW65616	Human bra
45	33	32.7	30	5	ABG47255	Human pep

ALIGNMENTS

RESULT 1
AAW42925
ID AAW42925 standard; peptide; 20 AA.
XX
AC AAW42925;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1318.
XX
KW Immunogenic peptide; immunogenic epitope; p2A protein; immune response;
KW antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006891.
XX
PR 19-APR-1996; 96US-0015644P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR Aar15517 BMP-8 pep
XX
XX Aaw42926 Immunogen
XX Aab69443 Synthetic
PT Aaw42924 Immunogen
PT Aab69441 Synthetic
PT Aab50903 Human sec
XX Abo45160 Novel hum
XX Abo26640 Protein a
PS Aab13951 Antiviral
XX Aau08313 Human his
CC Aab13952 Antiviral
CC Aau08314 Human his
CC Aab38287 Human sec
CC Aab38291 Gene 7 hu
CC Aab39297 Gene 43 h
CC Aab38184 Gene 9 hu
CC Aab51822 Gene 44 h
CC Aab75310 Gene 15 h
CC Aag62241 Human sec
XX Aab34679 Gene 44 h
XX Aab44387 Sequence

Claim 18; Page 112; 140pp; English.
Peptides AAW42922-30 are immunogenic peptides corresponding to immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are substantially similar to a portion of the amino acid sequence of the p2A protein of HAV corresponding to amino acids 792-980. The present peptide is derived from amino acids 823-842, and has a reactivity of 31.3% with acute sera. Compositions containing the peptides can be used to induce an immune response to HAV in a mammal. The peptides can also be used to detect the presence of antibodies against HAV in mammalian serum. The peptides can also be used to make an antibody against HAV by administering the peptide to a mammal
Sequence 20 AA;

Query Match 100.0%; Score 101; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 5,4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 VLPPPRKMKGLFSQAKISLF 20

RESULT 2
AA69442
ID AAB69442 standard; peptide; 21 AA.

AC AAB69442;

XX 20-APR-2001 (first entry)

DE Synthetic HAV P2A peptide, SEQ ID NO: 42.

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.

OS Synthetic.

XX WO200105824-A2.

XX 25-JAN-2001.

XX 14-JUL-2000; 2000WO-US019267.

XX 15-JUL-1999; 99US-0144412P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 2001-112681/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.

PS Claim 13; Page 95; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the IGM antibody reactivity

XX Sequence 21 AA;

Query Match 100.0%; Score 101; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 5,7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 VLPPPRKMKGLFSQAKISLF 20

RESULT 3

AAW42930

ID AAW42930 standard; peptide; 25 AA.

AC AAW42930;

XX 28-APR-1998 (first entry)

XX Immunogenic Hepatitis A virus peptide YK-1665.

XX Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.

OS Synthetic.

OS Hepatitis A virus.

XX WO9740147-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-US006891.

XX 19-APR-1996; 96US-0015644P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 1997-535831/49.

XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.

XX Claim 18; Page 112; 140pp; English.

XX Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. Compositions
CC containing the peptides can be used to induce an immune response to HAV
CC in a mammal. The peptides can also be used to detect the presence of
CC antibodies against HAV in mammalian serum. The peptides can also be used
CC to make an antibody against HAV by administering the peptide to a mammal

XX Sequence 25 AA;

Query Match 64.4%; Score 65; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLF 12
| | | | | | | | | | | | | |
Db 14 VLPPPRKMKGLF 25

RESULT 4

AAB69447

ID AAB69447 standard; peptide; 26 AA.

AC AAB69447;

XX 20-APR-2001 (first entry)

XX Synthetic HAV P2A peptide, SEQ ID NO: 47.

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.

OS Synthetic.

XX WO200105824-A2.

```

XX 25-JAN-2001.
PD
XX
XX PF
XX 14-JUL-2000; 2000WO-US019267.
XX
XX PR
XX 15-JUL-1999; 99US-0144412P.
XX
XX PA
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI
XX Fields HA, Khudyakov YE;
XX
XX DR
XX WPI; 2001-112681/12.
XX
XX PT
XX Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines.
XX
XX PS
XX Claim 13; Page 98; 130pp; English.
XX
XX CC
XX The present sequence is one of a number of synthetic peptides which are
XX immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
XX comprise antigenic epitopes of the major structural capsid polypeptides
XX or non-structural polypeptides of HAV with one or more glutamine
XX molecules at the carboxy end of the peptide. The peptides are used to
XX detect the presence of antibodies against HAV in mammalian serum, to
XX detect the presence of HAV in a human or animal through the binding of
XX the peptide to an antibody, to detect acute phase infection by detecting
XX IgM antibodies in mammalian serum and detecting convalescence in a
XX mammal. The peptides are used to detect or quantify HAV antibodies in
XX samples in clinical or research-based assays using immunoblotting,
XX fluorescence in situ hybridisation analysis, gel-mobility shift assays,
XX tracking of radioactive or bioluminescent markers, chromatography or
XX electrophoresis. The peptides are used to induce an immune response to
XX HAV when administered to a human or animal. Glutamine at the carboxy end
XX of the peptides enhances the IgM antibody reactivity
XX
XX SQ
XX Sequence 26 AA;
XX
XX Query Match 64.4%; Score 65; DB 4; Length 26;
XX Best Local Similarity 100.0%; Pred. No. 0.0038;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY
XX 1 VLPPPRKMKGLF 12
XX |||||
XX 14 VLPPPRKMKGLF 25
XX
XX RESULT 5
XX AAR15517
XX ID AAR15517 standard; protein; 23 AA.
XX
XX AC
XX AAR15517;
XX
XX CC
XX 25-MAR-2003 (revised)
XX DT 09-MAR-1992 (first entry)
XX
XX DE
XX BMP-8 peptide.
XX
XX KW
XX Cartilage; wound healing; tissue repair; BMP.
XX
XX OS
XX Bos taurus.
XX
XX PN
XX WO9118098-A.
XX
XX PD
XX 28-NOV-1991.
XX
XX PF
XX 16-MAY-1990; 90US-00525357.
XX
XX PR
XX 16-MAY-1990; 90US-00525357.
XX
XX PA
XX 15-JAN-1991; 91US-00641204.
XX
XX (GEMV ) GENETICS INST INC.
XX
XX PI
XX Hewick RM, Wang JH;

```

```

XX WPI; 1991-369252/50.
DR
DR N-PSDB; AAR15240, AAR15242, AAR15245, AAR15246.
XX
XX PT
XX New BMP-8 protein - useful in inducing cartilage and/or bone formation to
XX treat wounds and repair fractures and tissues, e.g. burns, incisions and
XX ulcers.
XX
XX PS
XX Claim 1(d); Page 46; 50pp; English.
XX
XX CC
XX This sequence shares some homology (i.e. Asn-Glu-Leu-Pro) with BMP-3 (see
XX WO88/00205 and WO89/10409). Pharmaceutical compsns. contg. BMP-8, which
XX comprises at least one of the fragments represented in AAR15517 and
XX AAR15522, can be used to aid bone and/or cartilage formation or wound
XX healing and tissue repair. The proteins are not very species specific so
XX can be used in domestic and farm animals as well as humans. See also
XX AAR15240-48, AAR15517 and AAR15522. (Updated on 25-MAR-2003 to correct PA
XX field.)
XX
XX SQ
XX Sequence 23 AA;
XX
XX Query Match 44.6%; Score 45; DB 2; Length 23;
XX Best Local Similarity 63.6%; Pred. No. 5;
XX Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY
XX 2 LPPPRKMKGLF 12
XX |||||
XX 4 LPPPNKLGIF 14
XX
XX RESULT 6
XX AAR42926
XX ID AAR42926 standard; peptide; 20 AA.
XX
XX AC
XX AAR42926;
XX
XX DT
XX 28-APR-1998 (first entry)
XX
XX DE
XX Immunogenic Hepatitis A virus peptide YK-1319.
XX
XX KW
XX Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.
XX
XX OS
XX Synthetic.
XX
XX PN
XX Hepatitis A virus.
XX
XX PD
XX WO9740147-A1.
XX
XX PF
XX 30-OCT-1997.
XX
XX PR
XX 18-APR-1997; 97WO-US006891.
XX
XX PR
XX 19-APR-1996; 96US-0015644P.
XX
XX PA
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI
XX Fields HA, Khudyakov YE;
XX
XX DR
XX WPI; 1997-535831/49.
XX
XX PT
XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
XX response to HAV in a mammal or to detect the presence of antibodies
XX against HAV in a mammal.
XX
XX PS
XX Claim 18; Page 112; 140pp; English.
XX
XX CC
XX Peptides AAR42922-30 are immunogenic peptides corresponding to
XX immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
XX substantially similar to a portion of the amino acid sequence of the P2A
XX protein of HAV corresponding to amino acids 792-980. The present peptide
XX is derived from amino acids 834-853, and has a reactivity of 27.1% with
XX acute sera. Compositions containing the peptides can be used to induce an
XX immune response to HAV in a mammal. The peptides can also be used to

```

CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal
 XX
 SQ Sequence 20 AA;

Query Match

Best Local Similarity 41.6%; Score 42; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FSQAKISLF 20

|||||
 1 FSQAKISLF 9

Db

RESULT 7

AAW42924

ID AAB69443 standard; peptide; 21 AA.

XX

AC AAB69443;

XX

DT 20-APR-2001 (first entry)

XX

DE Synthetic HAV P2A peptide, SEQ ID NO: 43.

XX

KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 antigen; major structural capsid polypeptide; HAV antibody detection.

XX

OS Hepatitis A virus.

XX

OS Synthetic.

XX

PN WO200105824-A2.

XX

PD 25-JAN-2001.

XX

PF 14-JUL-2000; 2000WO-US019267.

XX

PR 15-JUL-1999; 99US-0144412P.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Fields HA, Khudyakov YE;

XX

DR WPI; 2001-112681/12.

XX

XX Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.

XX

PS Claim 13; Page 95; 130pp; English.

XX

CC The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IgM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IgM antibody reactivity

XX Sequence 21 AA;

Query Match

Best Local Similarity 41.6%; Score 42; DB 4; Length 21;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FSQAKISLF 20

|||||
 1 FSQAKISLF 9

Db

RESULT 8

AAW42924

ID AAW42924 standard; peptide; 20 AA.

XX

AC AAW42924;

XX

DT 28-APR-1998 (first entry)

XX

DE Immunogenic Hepatitis A virus peptide YK-1317.

XX

KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
 antibody.

XX

OS Synthetic.

XX

OS Hepatitis A virus.

XX

PN WO9740147-A1.

XX

PD 30-OCT-1997.

XX

PF 18-APR-1997; 97WO-US006891.

XX

PR 19-APR-1996; 96US-0015644P.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Fields HA, Khudyakov YE;

XX

DR WPI; 1997-535831/49.

XX

PT Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.

XX

PS Claim 18; Page 112; 140pp; English.

XX

CC Peptides AAW42922-30 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the P2A
 CC protein of HAV corresponding to amino acids 792-980. The present peptide
 CC is derived from amino acids 810-829, and has a reactivity of 83.1% with
 CC acute sera. Compositions containing the peptides can be used to induce an
 CC immune response to HAV in a mammal. The peptides can also be used to
 CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal

XX Sequence 20 AA;

Query Match

Best Local Similarity 38.8%; Score 39; DB 2; Length 20;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRK 7

|||||
 14 VLPPPRK 20

Db

RESULT 9

AAB69441

ID AAB69441 standard; peptide; 21 AA.

XX

AC AAB69441;

XX

DT 20-APR-2001 (first entry)

XX

DE Synthetic HAV P2A peptide, SEQ ID NO: 41.

XX

KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 XX
 OS Hepatitis A virus.
 OS Synthetic.
 PN WO200105824-A2.
 XX
 PD 25-JAN-2001.
 XX
 XX 14-JUL-2000; 2000WO-US019267.
 PF
 XX 15-JUL-1999; 99US-0144412P.
 PR
 XX (USGS) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Fields HA, Khudyakov YE;
 PI WPI; 2001-112681/12.
 XX
 DR Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 PT
 XX Claim 13; Page 94; 130pp; English.
 PS
 XX The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IgM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IgM antibody reactivity
 XX
 SQ Sequence 21 AA;
 Query Match 38.6%; Score 39; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRK 7
 DB 14 VLPPPRK 20
 RESULT 10
 ABB50903
 ID ABB50903 standard; protein; 27 AA.
 XX
 AC ABB50903;
 XX
 DT 07-FEB-2002 (first entry)
 XX
 XX Human secreted protein encoded by gene 90 SEQ ID NO:856.
 DE
 XX Human; secreted protein; immunomodulatory; antisclerotic; anti-HIV;
 KW dermatological; immunosuppressive; antiinflammatory; immunostimulant;
 KW cytotatic; cardiant; vascular; anti-angiogenic; ophthalmological;
 KW neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnerary;
 KW antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
 KW multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
 KW human immunodeficiency virus; hyperproliferative disorder; wound healing;
 KW Chaga's disease; cardiovascular disease; Schmitz syndrome; chemotaxis;
 KW Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder;
 KW corneal graft neovascularisation; diabetic retinopathy; regeneration;

KW neurological disorder; Huntington's chorea; Alzheimer's disease;
 KW Parkinson's disease; infectious disease; chromosome 9.
 XX
 OS Homo sapiens.
 PN WO200162891-A2.
 XX
 PD 30-AUG-2001.
 XX
 XX 21-FEB-2001; 2001WO-US005614.
 PF
 XX 24-FEB-2000; 2000US-0184836P.
 PR
 XX 29-MAR-2000; 2000US-0193170P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ni J, Ebner R, Lafleur DW, Moore PA, Olsen HS, Rosen CA;
 PI Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
 PI Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrie AM, Fan P;
 PI Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G, Zeng Z;
 PI Greene JM;
 XX WPI; 2001-625724/72.
 DR
 XX Nucleic acids encoding 207 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating, e.g. cancers, Parkinson's disease
 PT and diabetic retinopathy.
 PT
 PS Disclosure; Page 210; 1533pp; English.
 XX
 CC ABB50301 to ABB51287 and ABA83194 to ABA83441 represent human secreted
 CC proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
 CC activities based on the tissues and cells the genes are expressed in.
 CC Example of these activities include: immunomodulatory; antisclerotic;
 CC dermatological; immunosuppressive; antiinflammatory; immunostimulant;
 CC anti-HIV; cytostatic; cardiant; anti-angiogenic; ophthalmological;
 CC neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;
 CC antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used
 CC in gene therapy and vaccine production. (I) and (II) can be used in the
 CC prevention, diagnosis and treatment of immune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus and human immunodeficiency virus
 CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
 CC Gaucher's disease), cardiovascular diseases (e.g. Schmitz syndrome,
 CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
 CC disorders (e.g. corneal graft neovascularisation and diabetic
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
 CC for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
 CC ABA83193 and ABB50300 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 27 AA;
 Query Match 36.6%; Score 37; DB 4; Length 27;
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LPPPRKMKGLFS 13
 DB 16 VPPPLKMPGVLA 27
 RESULT 11
 ABO45160
 ID ABO45160 standard; protein; 27 AA.
 XX
 AC ABO45160;
 XX
 DT 02-OCT-2003 (first entry)
 XX
 XX Novel human secreted protein #90 fragment #11.
 DE
 XX Human; gene therapy; autoimmune disorder; multiple sclerosis; cancer;

KW systemic lupus erythematosus; haematopoietic cell disorder; allergy;
KW agammaglobulinaemia; ataxia telangiectasia; blood coagulation disorder;
KW afibrinogenemia; thrombocytopenia; graft-versus-host disease; arthritis;
KW inflammatory condition; ischaemia-reperfusion injury; infectious disease;
KW hyperproliferative disorder; purpura; viral infection; regeneration;
KW bacterial infection; ulcer; Alzheimer's disease.
XX

OS Homo sapiens.

XX US2003045160-A1.

XX PD 03-APR-2003.

XX 07-DEC-2001; 2001US-00004860.

XX 06-JUN-1997; 97US-0048875P.

PR 06-JUN-1997; 97US-0048876P.

PR 06-JUN-1997; 97US-0048877P.

PR 06-JUN-1997; 97US-0048878P.

PR 06-JUN-1997; 97US-0048880P.

PR 06-JUN-1997; 97US-0048881P.

PR 06-JUN-1997; 97US-0048882P.

PR 06-JUN-1997; 97US-0048883P.

PR 06-JUN-1997; 97US-0048884P.

PR 06-JUN-1997; 97US-0048885P.

PR 06-JUN-1997; 97US-0048892P.

PR 06-JUN-1997; 97US-0048893P.

PR 06-JUN-1997; 97US-0048894P.

PR 06-JUN-1997; 97US-0048895P.

PR 06-JUN-1997; 97US-0048896P.

PR 06-JUN-1997; 97US-0048897P.

PR 06-JUN-1997; 97US-0048898P.

PR 06-JUN-1997; 97US-0048899P.

PR 06-JUN-1997; 97US-0048900P.

PR 06-JUN-1997; 97US-0048901P.

PR 06-JUN-1997; 97US-0048915P.

PR 06-JUN-1997; 97US-0048916P.

PR 06-JUN-1997; 97US-0048917P.

PR 06-JUN-1997; 97US-0048949P.

PR 06-JUN-1997; 97US-0048962P.

PR 06-JUN-1997; 97US-0048963P.

PR 06-JUN-1997; 97US-0048964P.

PR 06-JUN-1997; 97US-0048970P.

PR 06-JUN-1997; 97US-0048971P.

PR 06-JUN-1997; 97US-0048972P.

PR 06-JUN-1997; 97US-0048974P.

PR 06-JUN-1997; 97US-0049019P.

PR 06-JUN-1997; 97US-0049020P.

PR 06-JUN-1997; 97US-0049373P.

PR 06-JUN-1997; 97US-0049374P.

PR 06-JUN-1997; 97US-0049375P.

PR 06-JUN-1997; 97US-0057584P.

PR 06-JUN-1997; 97US-0057627P.

PR 06-JUN-1997; 97US-0057628P.

PR 06-JUN-1997; 97US-0057629P.

PR 06-JUN-1997; 97US-0057634P.

PR 06-JUN-1997; 97US-0057635P.

PR 06-JUN-1997; 97US-0057642P.

PR 06-JUN-1997; 97US-0057643P.

PR 06-JUN-1997; 97US-0057644P.

PR 06-JUN-1997; 97US-0057645P.

PR 06-JUN-1997; 97US-0057646P.

PR 06-JUN-1997; 97US-0057647P.

PR 06-JUN-1997; 97US-0057648P.

PR 06-JUN-1997; 97US-0057649P.

PR 06-JUN-1997; 97US-0057650P.

PR 06-JUN-1997; 97US-0057651P.

PR 06-JUN-1997; 97US-0057654P.

PR 06-JUN-1997; 97US-0057661P.

PR 06-JUN-1997; 97US-0057662P.

PR 06-JUN-1997; 97US-0057666P.

PR 06-JUN-1997; 97US-0057667P.

PR 06-JUN-1997; 97US-0057668P.

PR 05-SEP-1997; 97US-0057760P.

PR 05-SEP-1997; 97US-0057761P.

PR 05-SEP-1997; 97US-0057762P.

PR 05-SEP-1997; 97US-0057763P.

PR 05-SEP-1997; 97US-0057764P.

PR 05-SEP-1997; 97US-0057765P.

PR 05-SEP-1997; 97US-0057769P.

PR 05-SEP-1997; 97US-0057770P.

PR 05-SEP-1997; 97US-0057771P.

PR 05-SEP-1997; 97US-0057774P.

PR 05-SEP-1997; 97US-0057775P.

PR 05-SEP-1997; 97US-0057776P.

PR 05-SEP-1997; 97US-0057777P.

PR 05-SEP-1997; 97US-0057778P.

PR 18-DEC-1997; 97US-0070923P.

PR 04-JUN-1998; 98WO-US011422.

PR 15-JUL-1998; 98US-0092921P.

PR 30-JUL-1998; 98US-0094657P.

PR 04-DEC-1998; 98US-00205258.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;

PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;

PI Florence K, Laflleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;

PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;

PI Carter KC;

XX WPI; 2003-540804/51.

XX New isolated protein, useful for preparing a composition for diagnosing

PT or treating cancer, inflammatory, immune or infectious diseases.

PT New isolated protein, useful for preparing a composition for diagnosing

PT or treating cancer, inflammatory, immune or infectious diseases.

XX Disclosure; Page 57; 172pp; English.

XX The invention relates to an isolated HEMA80 protein. The protein is

CC useful for preparing a composition for diagnosing or treating autoimmune

CC disorders e.g. multiple sclerosis and systemic lupus erythematosus;

CC haematopoietic cell disorders e.g. agammaglobulinaemia and ataxia

CC telangiectasia; blood coagulation disorders e.g. afibrinogenemia and

CC thrombocytopenia; allergy; graft-versus-host disease; inflammatory

CC conditions e.g. ischaemia-reperfusion injury and arthritis;

CC hyperproliferative disorders e.g. cancer and purpura; infectious disease

CC e.g. viral infection and bacterial infection. The polynucleotide or

CC protein can be used to regenerate damaged tissue e.g. ulcers and

CC Alzheimer's disease. The present sequence represents the amino acid

CC sequence of a novel human secreted protein fragment. Note: The sequence

CC data for this patent did not form part of the printed specification but

CC was obtained in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?DocID=20030065160

XX Sequence 27 AA;

Query Match 36.8%; Score 37; DB 6; Length 27;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPPRKMKGLFS 13

Db 16 VPPPLKMPGVLA 27

RESULT 12

ABO26640

ID ABO26640 standard; protein; 27 AA.

XX ABO26640;

XX 10-SEP-2003 (first entry)

XX Protein associated with novel secreted protein gene 90 #11.

DE Secreted protein; precerebellin-like protein; sepsis; acne; psoriasis;

XX

KW neurodegenerative disorder; behavioural disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia; mania;
KW dementia; paranoia; psychosis; autism; immune disorder; infection;
KW inflammation; allergy; liver disorder; hepatoblastoma; jaundice;
KW hepatitis; immunological disorder; AIDS; leukaemia; rheumatoid arthritis;
XX
XX
OS Unidentified.
XX
XX US6525174-B1.
XX
XX
XX 25-FEB-2003.
XX
XX 04-DEC-1998; 98US-00205258.
XX
XX 06-JUN-1997; 97US-0048875P.
XX 06-JUN-1997; 97US-0048876P.
XX 06-JUN-1997; 97US-0048877P.
XX 06-JUN-1997; 97US-0048878P.
XX 06-JUN-1997; 97US-0048880P.
XX 06-JUN-1997; 97US-0048881P.
XX 06-JUN-1997; 97US-0048882P.
XX 06-JUN-1997; 97US-0048883P.
XX 06-JUN-1997; 97US-0048884P.
XX 06-JUN-1997; 97US-0048885P.
XX 06-JUN-1997; 97US-0048892P.
XX 06-JUN-1997; 97US-0048893P.
XX 06-JUN-1997; 97US-0048894P.
XX 06-JUN-1997; 97US-0048895P.
XX 06-JUN-1997; 97US-0048896P.
XX 06-JUN-1997; 97US-0048897P.
XX 06-JUN-1997; 97US-0048898P.
XX 06-JUN-1997; 97US-0048899P.
XX 06-JUN-1997; 97US-0048900P.
XX 06-JUN-1997; 97US-0048901P.
XX 06-JUN-1997; 97US-0048915P.
XX 06-JUN-1997; 97US-0048916P.
XX 06-JUN-1997; 97US-0048917P.
XX 06-JUN-1997; 97US-0048949P.
XX 06-JUN-1997; 97US-0048962P.
XX 06-JUN-1997; 97US-0048963P.
XX 06-JUN-1997; 97US-0048964P.
XX 06-JUN-1997; 97US-0048970P.
XX 06-JUN-1997; 97US-0048971P.
XX 06-JUN-1997; 97US-0048972P.
XX 06-JUN-1997; 97US-0048974P.
XX 06-JUN-1997; 97US-0049019P.
XX 06-JUN-1997; 97US-0049020P.
XX 06-JUN-1997; 97US-0049373P.
XX 06-JUN-1997; 97US-0049374P.
XX 06-JUN-1997; 97US-0049375P.
XX 05-SEP-1997; 97US-0057584P.
XX 05-SEP-1997; 97US-0057627P.
XX 05-SEP-1997; 97US-0057628P.
XX 05-SEP-1997; 97US-0057629P.
XX 05-SEP-1997; 97US-0057634P.
XX 05-SEP-1997; 97US-0057635P.
XX 05-SEP-1997; 97US-0057642P.
XX 05-SEP-1997; 97US-0057643P.
XX 05-SEP-1997; 97US-0057644P.
XX 05-SEP-1997; 97US-0057645P.
XX 05-SEP-1997; 97US-0057646P.
XX 05-SEP-1997; 97US-0057647P.
XX 05-SEP-1997; 97US-0057648P.
XX 05-SEP-1997; 97US-0057649P.
XX 05-SEP-1997; 97US-0057650P.
XX 05-SEP-1997; 97US-0057651P.
XX 05-SEP-1997; 97US-0057654P.
XX 05-SEP-1997; 97US-0057661P.
XX 05-SEP-1997; 97US-0057662P.
XX 05-SEP-1997; 97US-0057666P.
XX 05-SEP-1997; 97US-0057667P.
XX 05-SEP-1997; 97US-0057668P.

PR 05-SEP-1997; 97US-0057760P.
PR 05-SEP-1997; 97US-0057761P.
PR 05-SEP-1997; 97US-0057762P.
PR 05-SEP-1997; 97US-0057763P.
PR 05-SEP-1997; 97US-0057764P.
PR 05-SEP-1997; 97US-0057765P.
PR 05-SEP-1997; 97US-0057766P.
PR 05-SEP-1997; 97US-0057767P.
PR 05-SEP-1997; 97US-0057770P.
PR 05-SEP-1997; 97US-0057771P.
PR 05-SEP-1997; 97US-0057774P.
PR 05-SEP-1997; 97US-0057775P.
PR 05-SEP-1997; 97US-0057776P.
PR 05-SEP-1997; 97US-0057777P.
PR 05-SEP-1997; 97US-0057778P.
PR 18-DEC-1997; 97US-0070923P.
PR 04-JUN-1998; 98WO-US011422.
PR 15-JUL-1998; 98US-0092921P.
PR 30-JUL-1998; 98US-0094657P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Young P, Greene JM, Ferrie AM, Ruben SM, Rosen CA, Hu J;
PI Olsen HS, Ebner R, Brewer LA, Moore PA, Shi Y, Florence C;
PI Florence K, Lafleur DW, Ni J, Fan P, Wei Y, Fischer CL, Soppet DR;
PI Li Y, Zeng Z, Kyaw H, Yu G, Feng P, Dillon PJ, Endress GA;
PI Carter KC;
XX
DR WPI; 2003-511926/48.
XX
XX New precerebellin-like protein, useful for diagnosing or treating
PT neurodegenerative and behavioral disorders, immune disorders, liver
PT disorders, and cancer.
XX
PS Disclosure; Col 104; 156pp; English.
XX
CC The invention relates to an isolated protein comprising amino acid
CC residues 33-205 or 1-205 of a novel human secreted protein appearing as
CC ABO26252. The protein is encoded by one of 238 disclosed cDNA sequences
CC encoding 238 secreted proteins. ABO26252 is a precerebellin-like protein.
CC Also included are a composition comprising the protein and a carrier and
CC an isolated protein produced by expressing the protein cited above by a
CC cell, and recovering the protein. The proteins are useful for diagnosing
CC or treating neurodegenerative and behavioural disorders (e.g. Alzheimer's
CC disease, Parkinson's disease, Huntington's disease, schizophrenia, mania,
CC dementia, paranoia, psychoses or autism), immune disorders (e.g.
CC infection, inflammation, allergy), liver disorders (e.g. hepatoblastoma,
CC jaundice, hepatitis), immunological disorders (e.g. AIDS, leukaemia,
CC rheumatoid arthritis, sepsis, acne, psoriasis) and cancer. The present
CC sequence is a protein associated with one of the 238 disclosed novel
CC secreted proteins
XX
SQ Sequence 27 AA;
Query Match 36.6%; Score 37; DB 7; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 2 LPPPRKMKGLFS 13
:|||||: :
Db 16 VPPLKMPGVLA 27
RESULT 13
AAB13951
ID AAB13951 standard; peptide; 28 AA.
XX
AC AAB13951;
XX
DT 16-NOV-2000 (first entry)
XX
DE Antiviral peptide #17.
XX
KW Antiviral; antiulcer; human immunodeficiency virus; HIV;

KW herpes simplex virus; HSV; cold sore; aphthous ulcer;
 KW viral bronchial infection.

XX Synthetic.

XX Key Location/Qualifiers
 XX Modified-site 28
 XX /note= "C-terminal amide"

XX WO200032629-A2.

XX 08-JUN-2000.

XX 01-DEC-1999; 99WO-NL000732.

XX 01-DEC-1998; 98NL-01010692.

XX (TEWE-) STICHTING TECH WETENSCHAPPEN.

XX Van Nieuw Amerongen A, Veerman ECI, Van 'thof W, Nibbering PH;

XX WPI; 2000-412289/35.

XX Antiviral peptides comprising a domain of 10-25 amino acids, half of
 PT which is positively charged and half unchanged, useful for treatment of
 PT human immunodeficiency virus and herpes simplex virus.

XX Example 1; Page 7; 20pp; English.

XX The present invention relates to antiviral peptides containing a domain
 CC of 10-25 amino acids, where half the domain is positively charged and the
 CC other half is unchanged. The present sequence is an antiviral peptide of
 CC the invention. Oligomers consisting of at least two peptides coupled
 CC together are also included in the invention. The peptides are useful for
 CC treating viral infections such as human immunodeficiency virus (HIV) and
 CC herpes simplex virus (HSV) and for the treatment of cold sores, aphthous
 CC ulcers and viral bronchial infections

XX Sequence 28 AA;

Query Match 35.6%; Score 36; DB 3; Length 28;
 Best Local Similarity 57.1%; Pred. No. 1.7e+02;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 RKMKGFLFSQAKISL 19
 DB 12 RRYKRLFKELKFSL 25

RESULT 14

AAU08313
 ID AAU08313 standard; peptide; 29 AA.

XX AC AAU08313;

XX 04-DEC-2001 (first entry)

XX Human histatin 5 carboxy terminal based peptide 17.

XX Human; histatin; DHS; antimicrobial peptide; endocarditis;
 KW anti-infection; antiinflammatory; anti-repulsion; antifungal; antiviral.

XX Homo sapiens.

XX Synthetic.

XX Key Location/Qualifiers
 XX Modified-site 29
 XX /label= OTHER
 XX /note= "Iys is amidated"

XX WO200156627-A1.

XX 09-AUG-2001.

XX 12-JAN-2001; 2001WO-NL000019.

XX 12-JAN-2000; 2000NL-01014071.

XX (AMPH-) AM-PHARMA BV.

XX Van Nieuw Amerongen A, Veerman ECI, Van't Hof W;

XX WPI; 2001-514579/56.

XX Medical device, in particular implants and contact lenses for application
 PT into body of a patient, is coated with natural or synthetic antimicrobial
 PT peptides.

XX Disclosure; Page 7; 22pp; English.

XX The invention relates to medical device for application onto or into a
 CC body of a patient, coated with one or more naturally occurring peptides
 CC or proteins or synthetic peptides or their analogues having antimicrobial
 CC activity. The medical device is useful for application onto or into body
 CC of a patient and includes endotracheal tubes, intravenous catheters,
 CC urinary catheters, syringe needles, contact lenses, prosthetic implants,
 CC such as heart valves, bone implants, voice prostheses, pace makers, joint
 CC substitutes, dental implants and spinal implants. The medical device is
 CC also suitable for veterinary use. The peptide coated implants have anti-
 CC infection, antiinflammation and anti-repulsion activity, resulting in
 CC avoiding or inhibiting tissue loss after implantation and prevent bone
 CC degradation around the implants e.g. endocarditis. Cystatin-derived
 CC peptides inhibit both the proteolytic activity and growth of
 CC parapatogen, Porphyromonas gingivalis. Histatin and lactoferrin-derived
 CC peptides have increased antibacterial, antifungal and antiviral activity.
 CC By coating the medical device with an antimicrobial peptide or protein,
 CC the chance of the infection, including biofilm infection is effectively
 CC counteracted and the body of the patient is not oversensitive against
 CC such peptides or proteins. Resistant bacterial strain is not formed
 CC against such peptides. The present sequence is a synthetic oligomeric
 CC antimicrobial peptide based on human histatin 5 (DHS) carboxy terminus

XX Sequence 29 AA;

Query Match 35.6%; Score 36; DB 4; Length 29;
 Best Local Similarity 57.1%; Pred. No. 1.7e+02;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 RKMKGFLFSQAKISL 19
 DB 12 RRYKRLFKELKFSL 25

RESULT 15

AAAB13952

ID AAAB13952 standard; peptide; 28 AA.

XX AC AAAB13952;

XX 16-NOV-2000 (first entry)

XX Antiviral peptide #18.

XX Antiviral; antiulcer; human immunodeficiency virus; HIV;
 KW herpes simplex virus; HSV; cold sore; aphthous ulcer;
 KW viral bronchial infection.

XX Synthetic.

XX Key Location/Qualifiers
 XX Modified-site 28
 XX /note= "C-terminal amide"

XX WO200032629-A2.

XX 08-JUN-2000.


```

XX 01-DEC-1999; 99WO-NL000732.
XX 01-DEC-1998; 98NL-01010692.
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX Van Nieuw Amerongen A, Veerman ECI, Van 'thof W, Nibbering PH;
XX WPI; 2000-412289/35.
XX Antiviral peptides comprising a domain of 10-25 amino acids, half of
XX which is positively charged and half uncharged, useful for treatment of
XX human immunodeficiency virus and herpes simplex virus.
XX Example 1; Page 7; 20pp; English.
XX The present invention relates to antiviral peptides containing a domain
XX of 10-25 amino acids, where half the domain is positively charged and the
XX other half is uncharged. The present sequence is an antiviral peptide of
XX the invention. Oligomers consisting of at least two peptides coupled
XX together are also included in the invention. The peptides are useful for
XX treating viral infections such as human immunodeficiency virus (HIV) and
XX herpes simplex virus (HSV) and for the treatment of cold sores, aphthous
XX ulcers and viral bronchial infections
XX Sequence 28 AA;
XX
Query Match 34.7%; Score 35; DB 3; Length 28;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 6 RKMKGIFSQAKISL 19
Db 12 RYKRLFKLKFSL 25
Search completed: July 15, 2004, 14:46:41
Job time : 43.1333 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:40:36 ; Search time 8.53333 Seconds
(without alignments)
225.449 Million cell updates/sec

Title: US-09-171-432A-42
Perfect score: 101
Sequence: 1 VLPFRMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 6282

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	28.7	27	2	PI0029
2	28	27.7	14	2	C33098
3	28	27.7	20	4	I53672
4	28	27.7	22	1	MXKN2
5	28	27.7	27	2	C44524
6	27	26.7	12	2	B39690
7	27	26.7	20	2	A42267
8	27	26.7	24	2	PC2001
9	27	26.7	26	2	I49748
10	26	25.7	18	2	PC2280
11	26	25.7	21	2	A53733
12	26	25.7	25	2	S68005
13	26	25.7	27	2	US0676
14	26	25.7	28	2	JQ1035
15	25	24.8	12	2	PN0663
16	25	24.8	14	2	S65392
17	25	24.8	15	2	S67918
18	25	24.8	20	1	S06150
19	25	24.8	22	1	MXKN1
20	25	24.8	22	1	MXKN3
21	25	24.8	22	2	H83991
22	25	24.8	25	2	B44524
23	25	24.8	25	2	S07250
24	25	24.8	27	2	I49747
25	25	24.8	28	2	T14905
26	24	23.8	14	2	S14336
27	24	23.8	15	2	A26212
28	24	23.8	15	2	A60221
29	24	23.8	16	2	JH0517

30 24 23.8 20 2 B30208 hypothetical prote
31 24 23.8 25 2 A60842 somatostatin-25 -
32 24 23.8 26 2 C61590 beta-lactoglobulin
33 24 23.8 26 2 B60745 major outer membra
34 24 23.8 26 2 H83706 hypothetical prote
35 24 23.8 30 2 A22498 2-enoate reductase
36 23.5 23.3 26 2 A26070 neprilysin (EC 3.4
37 23 22.8 7 2 S71299 ICL2 protein - Par
38 23 22.8 14 2 PA0111 protein kinase (EC
39 23 22.8 14 2 S12904 tyrosine 3-monooxy
40 23 22.8 17 2 PN0587 Ig heavy chain DJ
41 23 22.8 19 2 PH1352 hypothetical prote
42 23 22.8 19 2 S25715 flavodoxin B - Azo
43 23 22.8 20 2 S17461 nicotinic acetylch
44 23 22.8 24 2 A27262 major pollen aller
45 23 22.8 24 2 D53288

ALIGNMENTS

RESULT 1

PL0029
Plasma protein Po2-S - pig (fragment)
N:Alternate names: alpha-1-glycoprotein homolog
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 11-May-2000
C:Accession: PL0029
R:Van de Weghe, A.; Coppieters, W.; Bauw, G.; Vandekerckhove, J.; Bouquet, Y.
Comp. Biochem. Physiol. B 90, 751-756, 1988
A:Title: The homology between the serum proteins Po2 in pig, Xk in horse and alpha 1B-gli
A:Reference number: PL0028; MUID:89250430; PMID:3248368
A:Accession: PL0029
A:Molecule type: protein
A:Residues: 1-27 <VAN>
A:Experimental source: serum
A:Note: 16-Pro was also found
C:Superfamily: alpha-1-B-glycoprotein; immunoglobulin homology
C:Keywords: glycoprotein; plasma
F:2/3/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 28.7%; Score 29; DB 2; Length 27;
Best Local Similarity 41.2%; Pred. No. 3.1e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 4; Gaps 1;

QY 3 PPRRMKGLFSQAKISL 19
||| |::: ||
Db 6 PPP---NLWAEAQTSLS 18

RESULT 2

C33098
223K exoantigen - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: plasmodium falciparum
C>Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000
C:Accession: C33098
R:Nichols, J.H.; Hager, L.P.
Submitted to the Protein Sequence Database, May 1990
A:Reference number: A33098
A:Accession: C33098
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-14 <NIC>

Query Match 27.7%; Score 28; DB 2; Length 14;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLPFRPKM 8
||| |:
Db 1 VLPPEKL 8

RESULT 3

I53672
somatotropin - synthetic
C;Species: synthetic
C;Date: 07-Jun-1996 #sequence_revision 31-Jul-1997 #text_change 19-May-2000
C;Accession: I53672
R;Bogolian, G.; Bilyeu, K.; O'Neil, J.P.
Gene 133, 17-22, 1993
A;Title: Genome rearrangements by residual IS10 elements in strains of *Escherichia coli*
A;Reference number: I53672; MUID:94040791; PMID:8224890
A;Accession: I53672
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-20 <BOG>
A;Cross-references: GB:S67119; NID:G455674; PIDN:AB28847.1; PID:G455675
A;Note: partial sequence of bovine somatotropin synthesized and expressed in *Escherichia*

Query Match 27.7%; Score 28; DB 4; Length 20;
Best Local Similarity 41.7%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 PPRKMKGLFSQA 15
| : ||| :
Db 3 PAMSLSLFANA 14

RESULT 4

MXN2
mu-conotoxin GIIIB [validated] - cone shell (*Conus geographus*)
N;Alternate names: geographutoxin II (GTx II); myotoxin II
C;Species: *Conus geographus* (geography cone)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 15-Sep-2000
C;Accession: A01787; B23579
R;Sato, S.; Nakamura, H.; Ohizumi, Y.; Kobayashi, J.; Hirata, Y.
FEBS Lett. 155, 277-280, 1983
A;Title: The amino acid sequences of homologous hydroxyproline-containing myotoxins from
A;Reference number: A91309; MUID:83210170; PMID:6852238
A;Accession: A01787
A;Molecule type: protein
A;Residues: 1-22 <SAT>
R;Gruz, L.J.; Gray, W.R.; Olivera, B.M.; Zeikus, R.D.; Kerr, L.; Yoshikami, D.; Moczyld
J. Biol. Chem. 260, 9280-9288, 1985
A;Title: *Conus geographus* toxins that discriminate between neuronal and muscle sodium ch
A;Reference number: A23579; MUID:85261316; PMID:2410412
A;Accession: B23579
A;Molecule type: protein
A;Residues: 1-22 <CRU>
R;Hall, J.M.; Alewood, P.F.; Craik, D.J.
submitted to the Brookhaven Protein Data Bank, April 1996
A;Reference number: A65705; PDB:1GIB
A;Contents: annotation; conformation by (1)H-NMR, residues 1-22
R;Hall, J.M.; Alewood, P.F.; Craik, D.J.
Biochemistry 35, 8824-8835, 1996
A;Title: Three-dimensional solution structure of mu-conotoxin GIIIB, a specific blocker
A;Reference number: A58590; MUID:96280640; PMID:8698418
A;Contents: annotation; conformation by (1)H-NMR
C;Superfamily: mu-conotoxin
C;Keywords: amidated carboxyl end; hydroxyproline; myotoxin; sodium channel inhibitor; v

F;15,4-20,10-21/Diulfide bonds: #status experimental
F;4,7/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;17/Modified site: 4-hydroxyproline (Pro) #status experimental
F;22/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 27.7%; Score 28; DB 1; Length 22;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PPRKMK 9
| : ||| :
Db 6 PPRKCK 11

RESULT 5

C44524

pregnancy-specific glycoprotein SBU-3-69 - sheep (fragment)
C;Species: *Ovis orientalis aries*, *Ovis ammon aries* (domestic sheep)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 12-Apr-1995
C;Accession: C44524
R;Atkinson, Y.H.
submitted to the Protein Sequence Database, June 1993
A;Reference number: A44524
A;Accession: C44524
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-27 <ATK>
C;Superfamily: pepsin
C;Keywords: glycoprotein

Query Match 27.7%; Score 28; DB 2; Length 27;
Best Local Similarity 37.5%; Pred. No. 4.5e+02;
Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PPRKMKGLFSQAQKISL 19
| : ||| :
Db 9 PLRNKDIYVGXITI 24

RESULT 6

B39690
neural cell adhesion molecule, cardiac splice form +, -, - rat (fragment)
C;Species: *Rattus norvegicus* (Norway rat)
C;Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 16-Jul-1999
C;Accession: B39690
R;Reyes, A.A.; Small, S.J.; Akesson, R.
Mol. Cell. Biol. 11, 1654-1661, 1991
A;Title: At least 27 alternatively spliced forms of the neural cell adhesion molecule m
A;Reference number: A39690; MUID:91141516; PMID:1996115
A;Accession: B39690
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-12 <REV>
A;Cross-references: GB:M63970
C;Keywords: cardiac muscle; cell adhesion; heart

Query Match 26.7%; Score 27; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PPRK 7
| : ||| :
Db 6 PPRE 10

RESULT 7

A42267
J-kappa recombination sequence-binding protein RBP-2N - mouse (fragment)
C;Species: *Mus musculus* (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C;Accession: A42267; S23798
R;Kawauchi, M.; Oka, C.; Shibayama, S.; Koromilas, A.E.; Matsunami, N.; Hamaguchi, Y.; F
J. Biol. Chem. 267, 4016-4022, 1992
A;Title: Genomic organization of mouse J-kappa recombination signal binding protein (RBP
A;Reference number: A42267; MUID:92156146; PMID:1740450
A;Accession: A42267
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-20 <KAW>
A;Cross-references: GB:M81866

Query Match 26.7%; Score 27; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 4.8e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 PPRKMK 8
| : ||| :
Db 13 PPRKRL 18

RESULT 8

PC2001
 major allergen - European chestnut (fragment)
 N/Alternate names: Cas s1
 C/Species: Castanea sativa (European chestnut)
 C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Mar-1999
 C/Accession: PC2001
 R/Kos, I.; Hoffmann-Sommergruber, K.; Ferreira, F.; Hirschwehr, R.; Ahorn, H.; Horak, F.
 Biochem. Biophys. Res. Commun. 196, 1086-1092, 1993
 A/Title: Purification, characterization and N-terminal amino acid sequence of a new major
 A/Reference number: PC2001; MUID:94071857; PMID:7504464
 A/Accession: PC2001
 A/Molecule type: protein
 A/Residues: 1-24 <KOS>
 C/Comment: This protein is antigenically related to the major birch pollen allergen Bet
 C/Superfamily: pathogenesis-related protein
 C/Keywords: pollen

Query Match 26.7%; Score 27; DB 2; Length 24;
 Best Local Similarity 55.6%; Pred. No. 5.8e+02;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY

1 VLPPrKMK 9

DB

12 VIPARLFK 20

RESULT 9

149748
 homeobox protein - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 15-Oct-1999
 C/Accession: I49748
 R/Dear, T.N.; Sanchez-Garcia, I.; Rabbitts, T.H.
 Proc. Natl. Acad. Sci. U.S.A. 90, 4431-4435, 1993
 A/Title: The HOX11 gene encodes a DNA-binding nuclear transcription factor belonging to
 A/Reference number: A47433; MUID:93281593; PMID:8099440
 A/Accession: I49748
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-26 <RES>
 A/Cross-references: GB:L08617; NID:g309311; PIDN:AAA37829.1; PID:g309312
 C/Superfamily: unassigned homeobox proteins; homeobox homology
 C/Keywords: DNA binding; homeobox; nucleus; transcription regulation

Query Match 26.7%; Score 27; DB 2; Length 26;
 Best Local Similarity 35.7%; Pred. No. 6.4e+02;
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY

3 PPRKMKGLFSQAK 16

DB

7 PPEKKHRANWSQLR 20

RESULT 10

PC2280
 prolylendopeptidase-inhibiting peptide - bovine (fragment)
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 28-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 09-May-1997
 C/Accession: PC2280
 R/Omori, T.; Nakagami, T.; Tanaka, H.; Maruyama, S.
 Biochem. Biophys. Res. Commun. 202, 809-815, 1994
 A/Title: Isolation of prolylendopeptidase-inhibiting peptides from bovine brain.
 A/Reference number: PC2280; MUID:94324971; PMID:8048952
 A/Accession: PC2280
 A/Molecule type: protein
 A/Residues: 1-18 <OHM>
 A/Experimental source: brain
 C/Superfamily: cytoskeletal keratin

Query Match 25.7%; Score 26; DB 2; Length 18;
 Best Local Similarity 35.7%; Pred. No. 9e+02;
 Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY

4 PPRKMKGLFSQAKISL 19

DB

8 PEVKMGLEGSDKLT 23

Best Local Similarity 42.9%; Pred. No. 6.2e+02;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY

2 LPPPRKMKGLFSQA 15

DB

1 MPPPLPARVDFSLA 14

RESULT 11

AS3733
 acetylcholine receptor alpha3 - rat (fragment)
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
 C/Accession: AS3733
 R/Yang, X.; McDonough, J.; Fyodorov, D.; Morris, M.; Wang, F.; Deneris, E.S.
 J. Biol. Chem. 269, 10252-10264, 1994
 A/Title: Characterization of an acetylcholine receptor alpha3 gene promoter and its acti
 A/Reference number: AS3733; MUID:94193711; PMID:8144606
 A/Accession: AS3733
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-21 <RES>
 A/Cross-references: EMBL:U04961; NID:g484068; PIDN:AAA18001.1; PID:g484069
 C/Superfamily: acetylcholine receptor
 C/Keywords: neurotransmitter receptor

Query Match

25.7%; Score 26; DB 2; Length 21;

Best Local Similarity

50.0%; Pred. No. 7.4e+02;

Matches

4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY

1 VLPPrKMK 8

DB

5 LVPPPLSM 12

RESULT 12

S68005
 hucolin, 35K chain - human (fragment)
 N/Alternate names: 36K HLA-cross-reactive plasma protein
 C/Species: Homo sapiens (man)
 C/Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 18-Feb-2000
 C/Accession: S68005; A38648
 R/Edgar, P.F.
 FEBS Lett. 375, 159-161, 1995
 A/Title: Hucolin, a new corticosteroid-binding protein from human plasma with structural
 A/Reference number: S68004; MUID:96087107; PMID:7498469
 A/Accession: S68005
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 1-25 <EDG>
 R/Haga, J.A.; She, J.X.; Kao, K.J.
 J. Biol. Chem. 266, 3695-3701, 1991
 A/Title: Biochemical characterization of 39-kDa class I histocompatibility antigen in pl

Query Match 25.7%; Score 26; DB 2; Length 25;
 Best Local Similarity 37.5%; Pred. No. 9e+02;
 Matches 6; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY

4 PPRKMKGLFSQAKISL 19

DB

8 PEVKMGLEGSDKLT 23

RESULT 13

JS0676
 hypothetical 3.1K protein (csk 5' region) - human
 C/Species: Homo sapiens (man)
 C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Mar-1995

C;Accession: JS0676
 R;Brauninger, A.; Holtrich, U.; Strebhardt, K.; Ruebsamen-Waigmann, H.
 Gene 110, 205-211, 1992
 A;Title: Isolation and characterization of a human gene that encodes a new subclass of P
 A;Reference number: JH0559; MUID:92165060; PMID:1371489
 A;Accession: JS0676
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-27 <BRA>
 A;Cross-references: EMBL:X59932
 A;Experimental source: lung

Query Match 25.7%; Score 26; DB 2; Length 27;
 Best Local Similarity 44.4%; Pred. No. 9.8e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPPRKMG 10
 ||| |
 Db 1 MPAREWPG 9

RESULT 14

JQ1035
 hypothetical 3.2K protein (type I IGF1R 5' region) - human
 C;Species: Homo sapiens (man)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
 C;Accession: JQ1035
 R;Cooke, D.W.; Bankert, L.A.; Roberts Jr., C.T.; LeRoith, D.; Casella, S.J.
 Biochem. Biophys. Res. Commun. 177, 1113-1120, 1991
 A;Title: Analysis of the human type I insulin-like growth factor receptor promoter regio
 A;Reference number: PQ0159; MUID:91282751; PMID:1711844
 A;Accession: JQ1035
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-28 <COO>

Query Match 25.7%; Score 26; DB 2; Length 28;
 Best Local Similarity 57.1%; Pred. No. 1e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPRKMG 10
 ||| |
 Db 9 PPRASG 15

RESULT 15

PNQ663
 dytrophin-associated glycoprotein A3a-II - rabbit (fragment)
 C;Species: Oryctolagus cuniculus (domestic rabbit)
 C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 07-May-1999
 C;Accession: PN0663
 R;Yoshida, M.; Mizuno, Y.; Nonaka, I.; Ozawa, E.
 J. Biochem. 114, 634-639, 1993
 A;Title: A dystrophin-associated glycoprotein, A3a (one of 43DAG doublets), is retained
 A;Reference number: PN0662; MUID:94156881; PMID:8113213
 A;Accession: PN0663
 A;Molecule type: protein
 A;Residues: 1-12 <YOS>
 C;Comment: This protein is retained in Duchenne type muscular dystrophy muscle.
 C;Keywords: Glycoprotein; skeletal muscle

Query Match 24.8%; Score 25; DB 2; Length 12;
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPP 5
 ||| |
 Db 4 LPPP 7

Search completed: July 15, 2004, 14:52:10
 Job time : 9.53333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:01 ; Search time 5.3333 Seconds
(without alignments)
195.263 Million cell updates/sec

Title: US-09-171-432A-42
Perfect score: 101
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 2073

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	27.7	17	1 APID_BOMPA	P81464 bombus pasc
2	28	27.7	22	1 CXM2_CONGE	P01524 conus geogr
3	26	25.7	24	1 LPER_STRFR	P45440 streptomyce
4	25	24.8	20	1 PSAF_MAIZE	P13193 zea mays (m
5	25	24.8	22	1 CXM1_CONGE	P01523 conus geogr
6	25	24.8	22	1 CXM3_CONGE	P05482 conus geogr
7	25	24.8	22	1 HELN_HELVI	P83427 heliothis v
8	24	23.8	14	1 MAST_VESPA	P21654 vespa basal
9	24	23.8	15	1 CBPB_PROAT	P19628 protopteris
10	24	23.8	16	1 IBP4_PIG	P24854 sus scrofa
11	24	23.8	28	1 SMS2_ORONI	P81029 oreochromis
12	24	23.8	30	1 ZENR_CLOTY	P1887 clostridium
13	23.5	23.3	26	1 NEP_PIG	P19621 sus scrofa
14	23	22.8	21	1 PIL3_ECOLI	P13948 escherichia
15	23	22.8	24	1 ACHA_ELBEL	P09688 electrophor
16	23	22.8	26	1 STP_BPT4	P18788 bacterioph
17	23	22.8	27	1 MULC_ECHML	P81797 echis multi
18	22	21.8	11	1 TKNA_HORSE	P01290 equus cabal
19	22	21.8	11	1 TKNA_RANRI	P29207 rana ridibu
20	22	21.8	15	1 PRP_MYCBO	P80149 mycobacteri
21	22	21.8	16	1 AU26_LITRA	P82393 litorea ran
22	22	21.8	22	1 BS23_SERPL	P83378 serratia pl
23	22	21.8	22	1 DBF1_SPIOL	P81572 spinacia ol
24	22	21.8	22	1 PSBQ_ORISA	P83646 oryza sativ
25	22	21.8	25	1 DBF3_SPIOL	P81570 spinacia ol
26	22	21.8	30	1 CALM_LYTPI	P05935 lytechinus
27	21.5	21.3	30	1 FMBB_BACNO	P17829 bacteroides
28	21	20.8	10	1 TMOF_AEADAE	P19425 aedes aegypt
29	21	20.8	11	1 BPP3_BOTIN	P30423 bothrops in
30	21	20.8	11	1 TKN3_PSEGU	P42988 pseudophryn
31	21	20.8	11	1 TKNA_CHKCK	P19850 gallus gall
32	21	20.8	11	1 TKNA_GADMO	P28498 gadus morhu
33	21	20.8	11	1 TKNA_RANCA	P22688 rana catesb

34	21	20.8	13	1 FIBB_HYLLA	P14472 hyllobates 1
35	21	20.8	13	1 TYI3_PHYKO	P04096 phyllomedus
36	21	20.8	14	1 MAST_VESXA	P01515 vespa xanth
37	21	20.8	15	1 ASP1_LACSN	P82648 lactobacill
38	21	20.8	15	1 UC06_MAIZE	P80612 zea mays (m
39	21	20.8	19	1 PSAE_CUCSA	P42047 cucumis eat
40	21	20.8	21	1 HCY6_MAISQ	P82308 maia squina
41	21	20.8	23	1 IBP2_PIG	P24853 sus scrofa
42	21	20.8	24	1 PCL1_PACGO	P82421 pachycondyl
43	21	20.8	24	1 PCL2_PACGO	P82422 pachycondyl
44	21	20.8	25	1 IPYR_PSEAN	P80898 pseudanabae
45	21	20.8	25	1 PCW1_PACGO	P82423 pachycondyl

ALIGNMENTS

RESULT 1					
APID_BOMPA					
ID	APID_BOMPA	STANDARD;	PRT;	17 AA.	
AC	P81464;				
DT	15-DEC-1998 (Rel. 37, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Apidaecin.				
OS	Bombus pascuorum (Brown bumble bee).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;				
OC	Apidae; Bombus.				
OX	NCBI_TaxID=65598;				
RN	[1]				
RP	SEQUENCE.				
RC	TISUE=Hemolymph;				
RX	MEDLINE=97362903; PubMed=9219367;				
RA	Rees J.A., Moniatte M., Bulet P.;				
RT	"Novel antibacterial peptides isolated from a European bumblebee,				
RT	Bombus pascuorum (Hymenoptera, Apoidea).";				
RL	Insect Biochem. Mol. Biol. 27:413-422(1997).				
CC	-!- FUNCTION: Antibacterial peptide active against Gram-negative				
CC	bacteria.				
CC	-!- INDUCTION: By bacterial infection.				
DR	InterPro: IPR004828; Apidaecin.				
DR	Pfam: PF00807; Apidaecin; 1.				
KW	Insect immunity; Antibiotic; Hemolymph.				
SQ	SEQUENCE 17 AA; 1963 MW; CDID0D02C8BC23D1 CRC64;				
Query Match		27.7%;	Score 28;	DB 1;	Length 17;
Best Local Similarity		80.0%;	Pred. No. 1.7e+02;		
Matches		4;	Conservative	1;	Mismatches 0;
				Indels	0;
				Gaps	0;
Qy	2 LPPPR 6				
	:				
Db	7 IPPPR 11				
RESULT 2					
CXM2_CONGE					
ID	CXM2_CONGE	STANDARD;	PRT;	22 AA.	
AC	P01524;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Mu-conotoxin GIIIB (Myotoxin II) (Geographutoxin II) (GTx-II).				
OS	Conus geographus (Geography cone).				
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;				
OC	Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;				
OC	Neogastropoda; Conoidea; Conidae; Conus.				
OX	NCBI_TaxID=6491;				
RN	[1]				
RP	SEQUENCE.				
RX	MEDLINE=85261316; PubMed=2410412;				
RA	Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,				
RA	Yoshikami D., Moczydlowski E.;				

RT "Conus geographus toxins that discriminate between neuronal and
 RT muscle sodium channels.",
 RL J. Biol. Chem. 260:9280-9288(1985).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=83210170; PubMed=6852238;
 RA Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;
 RT "The amino acid sequences of homologous hydroxyproline-containing
 RT myotoxins from the marine snail Conus geographus venom.";
 RL FEBS Lett. 155:277-280(1983).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=96280640; PubMed=8688418;
 RA Hill J.M., Alewood P.F., Craik D.J.;
 RT "Three-dimensional solution structure of mu-conotoxin GIIIB, a
 RT specific blocker of skeletal muscle sodium channels.";
 RL Biochemistry 35:8824-8835(1996).
 CC -!- FUNCTION: Mu-conotoxins bind and block voltage-sensitive sodium
 CC channel (VSSC).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS. MU-TYPE
 CC FAMILY.
 CC PIR; A01787; MXKN2.
 DR PDB; 1GIB; 08-NOV-96.
 DR InterPro; IPR008036; Mu-conotoxin.
 DR Pfam; PF05374; Mu-conotoxin; 1.
 DR Toxin; Sodium channel inhibitor; Hydroxylation; Amidation;
 KW 3D-structure.
 FT DISULFID 3 15
 FT DISULFID 4 20
 FT DISULFID 10 21
 FT MOD RES 6 6
 FT MOD RES 7 7
 FT MOD RES 17 17
 FT MOD RES 22 22
 FT MOD RES 22 22
 FT AMIDATION.
 SQ SEQUENCE 22 AA; 2599 MW; F50402BA93199E01 CRC64;
 Query Match 27.7%; Score 28; DB 1; Length 22;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 PPRKMK 9
 |||||
 DB 6 PPRKCK 11
 RESULT 3
 LPER STRFR STANDARD; PRT; 24 AA.
 AC P4540;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Brythromycin resistance leader peptide (23S rRNA methylase leader
 DE peptide).
 OS Streptomyces fradiae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1906;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88169508; PubMed=3127381;
 RA Kamimura S., Weisblum B.;
 RT "Translational attenuation control of ermSF, an inducible resistance
 RT determinant encoding rRNA N-methyltransferase from Streptomyces
 RT fradiae.";
 RL J. Bacteriol. 170:1800-1811(1988).
 CC -!- FUNCTION: THIS PEPTIDE IS INVOLVED IN THE CONTROL MECHANISM OF
 CC THE SYNTHESIS OF THE MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B
 CC RESISTANCE PROTEIN.
 CC -!- CAUTION: It is uncertain whether Met-1 or Met-3 is the initiator.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M19269; AAA26741.1; -;
 DR Antibiotic resistance; Leader peptide.
 SQ SEQUENCE 24 AA; 2529 MW; 756C1635F8C5F06A CRC64;
 Query Match 25.7%; Score 26; DB 1; Length 24;
 Best Local Similarity 80.0%; Pred. No. 5.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPP 5
 :|||
 DB 14 LLPPP 18
 RESULT 4
 PSFAF_MAIZE STANDARD; PRT; 20 AA.
 ID PSFAF_MAIZE
 AC P13193;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Photosystem I reaction centre subunit III (Light-harvesting complex I
 DE 17 kDa protein) (PSI-F) (Fragment).
 GN PSFAF.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. N273;
 RX MEDLINE=90033230; PubMed=2680596;
 RA Anandan S., Vainstein A., Thorner J.P.;
 RT "Correlation of some published amino acid sequences for photosystem I
 RT polypeptides to a 17 kDa LHCI pigment-protein and to subunits III and
 RT IV of the core complex.";
 RL FEBS Lett. 256:150-154(1989).
 CC -!- FUNCTION: Probably participates in efficiency of electron transfer
 CC from plastocyanin to P700 (or cytochrome c553 in algae and
 CC cyanobacteria). This plastocyanin-docking protein contributes to
 CC the specific association of plastocyanin to PSI.
 CC -!- SUBCELLULAR LOCATION: Associated with luminal side of the
 CC thylakoid membrane.
 CC -!- SIMILARITY: Belongs to the psaf family.
 DR PIR; S06150; S06150.
 DR MaizeDB; 69542; -;
 KW Photosynthesis; Photosystem I; Chloroplast; Thylakoid; Membrane.
 FT NON TER 20
 FT NON TER 20
 SQ SEQUENCE 20 AA; 2110 MW; F0499F98F6188997 CRC64;
 Query Match 24.8%; Score 25; DB 1; Length 20;
 Best Local Similarity 33.3%; Pred. No. 6.2e+02;
 Matches 5; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 2 LPPPRKMKGLFSQAK 16
 |||||
 DB 5 LTPPKSKAFKXEK 19
 RESULT 5
 CXMI CONGE STANDARD; PRT; 22 AA.
 ID CXMI_CONGE
 AC P01523;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mu-conotoxin GIIIA (Myotoxin I) (Geographutoxin I) (GTX-I).
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE.
RX MEDLINE=85261316; PubMed=2410412;
RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,
RA Yoshikami D., Moczydlowski E.;
RT "Conus geographus toxins that discriminate between neuronal and
RT muscle sodium channels."; 288(1985).
RL J. Biol. Chem. 260:9280-9288(1985).
RN [2]
RP SEQUENCE.
RX MEDLINE=83210170; PubMed=6852238;
RA Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;
RA "The amino acid sequences of homologous hydroxyproline-containing
RT myotoxins from the marine snail Conus geographus venom.";
RN FEBS Lett. 155:277-280(1983).
RN [3]
RP DISULFIDE BONDS.
RX MEDLINE=90249506; PubMed=2338142;
RA Hidaka Y., Sato K., Nakamura H., Kobayashi J., Ohizumi Y.,
RA Simonishi Y.;
RT "Disulfide pairings in geographutoxin I, a peptide neurotoxin from
RT Conus geographus.";
RN FEBS Lett. 264:29-32(1990).
RN [4]
RP STRUCTURE BY NMR.
RX MEDLINE=91122275; PubMed=1991506;
RA Ott K.-H., Becker S., Gordon R.D., Rueterjans H.;
RT "Solution structure of mu-conotoxin GIIIA analysed by 2D-NMR and
RT distance geometry calculations.";
RN FEBS Lett. 278:160-166(1991).
RN [5]
RP STRUCTURE BY NMR.
RX MEDLINE=91299744; PubMed=2069951;
RA Lancelin J.-M., Kohda D., Tate S.-I., Yanagawa Y., Abe T., Satake M.,
RA Inagaki F.;
RT "Tertiary structure of conotoxin GIIIA in aqueous solution.";
RN Biochemistry 30:6908-6916(1991).
CC -!- FUNCTION: Mu-conotoxins bind and block voltage-sensitive sodium
CC channel (VSSC).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS. MU-TYPE
CC FAMILY.
DR PIR; A01786; MXKN1.
DR PDB; 1TCG; 31-JAN-94.
DR PDB; 1TCH; 31-JAN-94.
DR PDB; 1TCJ; 31-JAN-94.
DR PDB; 1TCK; 31-JAN-94.
DR InterPro; IPR008036; Mu-conotoxin.
DR Pfam; PF05374; Mu-conotoxin; 1.
KW Toxin; Sodium channel inhibitor; Hydroxylation; Amidation;
KW 3D-structure.
FT DISULFID 3 15
FT DISULFID 4 20
FT DISULFID 10 21
FT MOD_RES 6 6 HYDROXYLATION.
FT MOD_RES 7 7 HYDROXYLATION.
FT MOD_RES 17 17 HYDROXYLATION.
FT MOD_RES 22 22 AMIDATION.
FT HELIX 13 16
FT TURN 19 21
SQ SEQUENCE 22 AA; 2568 MW; F6CB02ADB359813C CRC64;
Query Match 24.8%; Score 25; DB 1; Length 22;
Best Local Similarity 66.7%; Pred. No. 6.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PPRKMK 9
DB 6 PPKCK 11
RESULT 6
CXM3_CONGE STANDARD; PRT; 22 AA.
AC P05482;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mu-conotoxin GIIIC.
OS Conus geographus (Geography cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6491;
RN [1]
RP SEQUENCE.
RX MEDLINE=85261316; PubMed=2410412;
RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,
RA Yoshikami D., Moczydlowski E.;
RT "Conus geographus toxins that discriminate between neuronal and
RT muscle sodium channels."; 288(1985).
RL J. Biol. Chem. 260:9280-9288(1985).
CC -!- FUNCTION: Mu-conotoxins bind and block voltage-sensitive sodium
CC channel (VSSC).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS. MU-TYPE
CC FAMILY.
DR PIR; C23579; MXKN3.
DR HSSP; P01524; 1GIB.
DR InterPro; IPR008036; Mu-conotoxin.
DR Pfam; PF05374; Mu-conotoxin; 1.
KW Toxin; Sodium channel inhibitor; Hydroxylation; Amidation.
FT DISULFID 3 15
FT DISULFID 4 20 BY SIMILARITY.
FT DISULFID 10 21 BY SIMILARITY.
FT MOD_RES 6 6 HYDROXYLATION.
FT MOD_RES 7 7 HYDROXYLATION.
FT MOD_RES 17 17 HYDROXYLATION.
FT MOD_RES 22 22 AMIDATION.
SQ SEQUENCE 22 AA; 2553 MW; F50402BA92A9813C CRC64;
Query Match 24.8%; Score 25; DB 1; Length 22;
Best Local Similarity 66.7%; Pred. No. 6.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PPRKMK 9
DB 6 PPKCK 11
RESULT 7
HELN_HELVI STANDARD; PRT; 22 AA.
AC P83427;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Helicoin.
OS Heliothis virescens (Noctuid moth) (Owllet moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Noctuoidea;
OC Noctuidae; Heliothinae; Heliothinae.
OX NCBI_TaxID=7102;
RN [1]
RP SEQUENCE, FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, TISSUE SPECIFICITY.
RP MASS SPECTROMETRY, PYRROLIDONE CARBOXYLIC ACID, AND
RP CARBOHYDRATE-LINKAGE SITE THR-7.

RC TISSUE=Hemolymph;
 RA Bulet P., Lambert M., Charlet M., Sabatier L., Rabel D.;
 RL Submitted (JUL-2002) to Swiss-Prot.
 CC -!- FUNCTION: Has antibacterial activity, preferentially against Gram-
 CC negative bacteria.
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Hemolymph.
 CC -!- INDUCTION: By bacterial infection.
 CC -!- MASS SPECTROMETRY: MW=2901.89; METHOD=MALDI.
 CC -!- SIMILARITY: Belongs to the lebecin family.
 CC Insect immunity; Antibiotic; Hemolymph; Glycoprotein;
 KW Pyrrolidone carboxylic acid.
 FT MOD_RES 1 7 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 7 7 O-LINKED (GALNAc).
 FT SEQUENCE 22 AA; 2714 MW; 1F4368EB8817F059 CRC64;
 SQ
 Query Match 24.8%; Score 25; DB 1; Length 22;
 Best Local Similarity 30.8%; Pred. No. 6.9e+02;
 Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 3 PPRKMKGLFSQA 15
 DB 10 PPQPRRPVIMRA 22
 ||| : : :
 RESULT 8
 MAST_VESBA STANDARD; PRT; 14 AA.
 AC P21654;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mastoparan B.
 OS Vespa basalis (Hornet).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Vespoidea;
 OC Vespidae; Vespinae; Vespa.
 OC NCBI_TaxID=7444;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=91174755; PubMed=2006909;
 RA Lo C.-L., Huang L.-L.;
 RT "Structure and biological activities of a new mastoparan isolated
 RT from the venom of the hornet Vespa basalis.";
 RL Biochem. J. 274:453-456(1991).
 CC -!- FUNCTION: Mast cell degranulating peptide. Activates G proteins
 CC that couple to phospholipase C.
 DR PIR; S14336; S14336
 KW Mast cell degranulation; Amidation.
 FT MOD_RES 14 14 AMIDATION.
 FT SEQUENCE 14 AA; 1613 MW; D35944CA19A19A2 CRC64;
 SQ
 Query Match 23.8%; Score 24; DB 1; Length 14;
 Best Local Similarity 50.0%; Pred. No. 6.3e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 7 KMKGLFSQAK 16
 DB 2 KLSIVSWAK 11
 ||| : : :
 RESULT 9
 CBP_PROAT STANDARD; PRT; 15 AA.
 AC P19628;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carboxypeptidase B (EC 3.4.17.2) (Fragment).
 OS Prototermus aethiopicus (Marbled lungfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Dipnoi; Lepidosireniformes; Protopterygia; Protopterus.
 OX NCBI_TaxID=7886;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pancreas;
 RX MEDLINE=73025047; PubMed=5079891;
 RA Reeck G.R., Neurath H.;
 RT "Isolation and characterization of pancreatic procarboxypeptidase B
 RT and carboxypeptidase B of the African lungfish.";
 RL Biochemistry 11:3947-3955(1972).
 CC -!- CATALYTIC ACTIVITY: Peptidyl-L-lysine (or L-arginine) + H(2)O =
 CC peptide + L-lysine (or L-arginine).
 CC -!- SIMILARITY: Belongs to peptidase family M14.
 DR PIR; A26212; A26212.
 DR MEROPS; M14.003;
 DR InterPro; IPR000834; Peptidase_M14.
 DR PROSITE; PS00132; CARBOXYPEPT_ZN_1; PARTIAL.
 DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; PARTIAL.
 KW Hydrolase; Carboxypeptidase; Metalloprotease; Zinc; Zymogen.
 FT PROPEP 1 >15 ACTIVATION PEPTIDE.
 FT NON_TER 15 15
 FT SEQUENCE 15 AA; 1749 MW; 124C910D937BED65 CRC64;
 SQ
 Query Match 23.8%; Score 24; DB 1; Length 15;
 Best Local Similarity 50.0%; Pred. No. 6.7e+02;
 Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 PPRKMKMG 10
 DB 2 FTRPSFNG 9
 ||| : : :
 RESULT 10
 IBP4_PIG STANDARD; PRT; 16 AA.
 AC P24854;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Insulin-like growth factor binding protein 4 (IGFBP-4) (IBP-4)
 DE (IGF-binding protein 4) (fragment).
 GN IGFBP4.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92109718; PubMed=1722398;
 RA Coleman M.E., Pan Y.-C.E., Ehterton T.D.;
 RT "Identification and NH2-terminal amino acid sequence of three
 RT insulin-like growth factor-binding proteins in porcine serum.";
 RL Biochem. Biophys. Res. Commun. 181:1131-1136(1991).
 CC -!- FUNCTION: IGF-binding proteins prolong the half-life of the IGFs
 CC and have been shown to either inhibit or stimulate the growth
 CC promoting effects of the IGFs on cell culture. They alter the
 CC interaction of IGFs with their cell surface receptors.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Contains 1 IGFBP domain.
 CC -!- SIMILARITY: Contains 1 thyroglobulin type-I domain.
 DR PIR; JH0517; JH0517.
 DR InterPro; IPR000867; Insl_gro_fac_pr.
 DR InterPro; IPR000716; Thyroglobulin_1.
 DR PROSITE; PS00222; IGF_BINDING; PARTIAL.
 DR PROSITE; PS00484; THYROGLOBULIN_1; PARTIAL.
 KW Growth factor binding.
 FT NON_TER 16 16
 FT SEQUENCE 16 AA; 1799 MW; 409884009655E2 CRC64;
 SQ
 Query Match 23.8%; Score 24; DB 1; Length 16;
 Best Local Similarity 57.1%; Pred. No. 7.2e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
QY      3 PPRKWK 9
      |||||.
Db      7 PPFSEK 13

RESULT 11
SMS2_ORENI
ID_SMS2_ORENI STANDARD; PRT; 28 AA.
AC P81029;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin II precursor [Contains: [Tyr21,Gly24]somatostatin-28;
DE [Tyr7,Gly10]somatostatin-14] (Fragment)
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE.
RX MEDLINE=95384941; PubMed=7656183;
RA Nguyen T.M., Wright J.R. Jr., Nielsen P.F., Conlon J.M.;
RT "Characterization of the pancreatic hormones from the Brockmann body
RT of the tilapia: implications for islet xenograft studies.";
RL Comp. Biochem. Physiol. 111C:33-44(1995).
CC -!- FUNCTION: Somatostatin inhibits the release of somatotropin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the somatostatin family.
DR InterPro: IPR004250; Somatostatin.
KW Cleavage on pair of basic residues; Hormone; Multigene family.
FT NON_TER 1
FT PEPTIDE 1 28 [TYR21, GLY24] SOMATOSTATIN-28.
FT PEPTIDE 15 28 [TYR7, GLY10] SOMATOSTATIN-14.
FT DISULFID 17 28
FT SEQUENCE 28 AA; 3155 MW; 47C049F4866EF4AC CRC64;

Query Match 23.8%; Score 24; DB 1; Length 28;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 PPRKWK 9
      |||||.
Db      9 PPRKWK 14

RESULT 12
2ENR_CLOTY
ID_2ENR_CLOTY STANDARD; PRT; 30 AA.
AC P11887;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 2-enoate reductase [EC 1.3.1.31] (Fragment).
OS Clostridium tyrobutyricum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1519;
RN [1]
RP SEQUENCE.
RX MEDLINE=85225966; PubMed=4005048;
RA Kuno S., Bacher A., Simon H.;
RT "Structure of enoate reductase from a Clostridium tyrobutyricum (C.
RT spec. la1).";
RL Biol. Chem. Hoppe-Seyler 366:463-472(1985).
CC -!- FUNCTION: INVOLVED IN FERMENTATION OF AA (STICKLAND REACTION)
CC SUCH AS LEUCINE, ISOLEUCINE, VALINE AND PHENYLALANINE.
CC -!- CATALYTIC ACTIVITY: Butanoate + NAD(+) = 2-butenate + NADH.
CC -!- COFACTOR: Iron sulfur flavoprotein.
CC -!- SUBUNIT: Dodecamer; tetramer of trimers.
DR PIR; A22498; A22498.
```

```
KW Oxidoreductase; NAD; Flavoprotein; Iron-sulfur.
FT NON_TER 30
SQ SEQUENCE 30 AA; 3299 MW; AD96FB5BB2080920 CRC64;

Query Match 23.8%; Score 24; DB 1; Length 30;
Best Local Similarity 54.5%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      7 KMKGLFSQAKI 17
      |||||.
Db      2 KMKSLFEVIKI 12

RESULT 13
NEP_PIG
ID_NEP_PIG STANDARD; PRT; 26 AA.
AC P19621;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neprilysin (EC 3.4.24.11) (Neutral endopeptidase) (NEP)
DE (Enkephalinase) (Fragment).
GN MME.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=87156592; PubMed=3548708;
RA Fuchner I.S., Pappin D.J.C., Kenny A.J.;
RT "The N-terminal amino acid sequence of pig kidney endopeptidase-24.11
RT shows homology with pro-sucrase-isomaltase.";
RL Biochem. J. 240:305-308(1986).
CC -!- FUNCTION: Thermolysin-like specificity, but is almost confined on
CC acting on polypeptides of up to 30 amino acids. Biologically
CC important in the destruction of opioid peptides such as Met- and
CC leu-enkephalins by cleavage of a Gly-Phe bond.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage at the amino group of
CC hydrophobic residues in insulin, casein, hemoglobin, and a number
CC of other proteins and polypeptides.
CC -!- COFACTOR: Binds 1 zinc ion per subunit.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: Belongs to peptidase family M13.
DR MEROPS; M13.001; -.
DR InterPro: IPR006025; Pept M Zn BS.
DR PROSITE: PS00142; ZINC_PROTEASE; PARTIAL.
KW Hydrolase; Metalloprotease; Zinc; Transmembrane; Glycoprotein.
FT NON_TER 1
FT NON_TER 26
SQ SEQUENCE 26 AA; 3038 MW; 3848804A9DDF7DEF CRC64;

Query Match 23.3%; Score 23.5; DB 1; Length 26;
Best Local Similarity 41.2%; Pred. No. 1.4e+03;
Matches 7; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY      3 PPRKMKGLFSQAKISL 19
      |||||.
Db      1 PKPKK-KQRWTFLEISL 16

RESULT 14
PIL3_ECOLI
ID_PIL3_ECOLI STANDARD; PRT; 21 AA.
AC P13948;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Fimbrial protein precursor (Pilin) (Fragment).
GN TRAA.
OS Escherichia coli.
```

OG Plasmid ColB4-K98.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87008371; PubMed=3531163;
RA Finlay B.B., Frost L.S., Paranych W.;
RT "Origin of transfer of IncF plasmids and nucleotide sequences of the
RT type II oriT, traM, and traY alleles from ColB4-K98 and the type IV
RT traY allele from RI00-1.";
RL J. Bacteriol. 168:132-139(1986).
CC -I- FUNCTION: PROPILIN IS THE PRECURSOR OF THE SEXPILOS SUBUNIT. THE
CC SEXPILOS ARE FILAMENTOUS SURFACE APPENDAGES REQUIRED FOR CELL-TO-
CC CELL CONTACT DURING BACTERIAL CONJUGATION.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M15135; AAB04666.1; -
DR InterPro: IPR008873; TraA.
DR Pfam: PF05513; TraA; 1.
KW Plasmid; Conjugation; Fimbria.
FT PROPEP 1 >21
FT NON TER 21 21
SQ SEQUENCE 21 AA; 2198 MW; 076C0D0C5B9D14EA CRC64;

Query Match 22.8%; Score 23; DB 1; Length 21;
Best Local Similarity 55.6%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 PRKMKGLFS 13
| | | | |
Db 13 PVKKKSFSS 21

RESULT 15
ACUA ELEEL STANDARD; PRT; 24 AA.
ID ACHA_ELEEL
AC P09688;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetylcholine receptor protein, alpha chain (Fragment).
OS Electrophorus electricus (Electric eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Gymnotiformes;
OC Electrophoridae; Electrophorus.
OX NCBI_TaxID=8005;
RN [1]
RP SEQUENCE
RX MEDLINE=83065199; PubMed=6959131;
RA Conti-Tronconi B.M., Hunkapiller M.W., Lindstrom J.M., Raftery M.A.;
RT "Subunit structure of the acetylcholine receptor from Electrophorus
RT electricus.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6489-6493(1982).
CC -I- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -I- SUBUNIT: Pentamer of two alpha chains, and one each of the beta,
CC delta, and gamma chains.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: Belongs to the ligand-gated ionic channel family.
DR PIR; A27262; A27262.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006201; Neur_channel.

DR Pfam: PF02931; Neur_chan_LBD; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; PARTIAL.
KW Receptor; Postsynaptic membrane; Ionic channel; Transmembrane.
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2802 MW; 3FE05AFB6B4983AE CRC64;

Query Match 22.8%; Score 23; DB 1; Length 24;
Best Local Similarity 62.5%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 RRMKGLFS 13
| | | | |
Db 6 RLKKNLFS 13

Search completed: July 15, 2004, 14:47:31
Job time : 6.33333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:31 ; Search time 25.7333 Seconds
(without alignments)
245.221 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101

Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 18020

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25;*

1: sp_arChaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	34	33.7	18	4	Q16167	Q16167 homo sapien
2	33	32.7	29	11	Q8C837	Q8C837 mus musculus
3	32	31.7	16	6	Q9TQZ7	Q9TQZ7 bos taurus
4	32	31.7	27	12	Q37240	Q37240 hepatitis c
5	32	31.7	27	12	Q37239	Q37239 hepatitis c
6	32	31.7	27	12	Q37244	Q37244 hepatitis c
7	32	31.7	27	12	Q37242	Q37242 hepatitis c
8	32	31.7	27	12	Q37222	Q37222 hepatitis c
9	32	31.7	27	12	Q37241	Q37241 hepatitis c
10	32	31.7	27	12	Q37238	Q37238 hepatitis c
11	32	31.7	27	12	Q37237	Q37237 hepatitis c
12	32	31.7	27	12	Q37243	Q37243 hepatitis c
13	32	31.7	27	12	Q37223	Q37223 hepatitis c
14	31	30.7	24	4	Q8TE34	Q8TE34 homo sapien
15	31	30.7	30	4	Q9BZF8	Q9BZF8 homo sapien
16	31	30.7	30	16	Q7VFY7	Q7VFY7 helicobacte

17	30	29.7	21	4	Q9UCC3	Q9UCC3 homo sapien
18	30	29.7	22	6	Q9TNN1	Q9TNN1 monodelphis
19	30	29.7	23	13	Q9PV93	Q9PV93 gallus gall
20	29	28.7	20	2	Q83VI3	Q83VI3 vibrio chol
21	29	28.7	27	12	Q37236	Q37236 hepatitis c
22	28	27.7	15	10	P82439	P82439 nicotiana t
23	28	27.7	20	2	Q53370	Q53370 escherichia
24	28	27.7	27	6	Q9TRE2	Q9TRE2 ovis aries
25	28	27.7	27	8	Q9TM43	Q9TM43 cyanidium c
26	28	27.7	30	16	Q8EB19	Q8EB19 shewanella
27	27.5	27.2	23	13	Q9PRX7	Q9PRX7 xenopus lae
28	27	26.7	14	11	Q99PB8	Q99PB8 mus musculu
29	27	26.7	18	4	Q3UCT9	Q3UCT9 homo sapien
30	27	26.7	19	2	Q9RBV1	Q9RBV1 pseudomonas
31	27	26.7	20	12	Q85617	Q85617 reovirus (t
32	27	26.7	22	12	Q9WH75	Q9WH75 chandipura
33	27	26.7	22	12	Q85664	Q85664 reovirus (t
34	27	26.7	24	10	Q9S9Q4	Q9S9Q4 castanea sa
35	27	26.7	26	2	Q9F4H7	Q9F4H7 klebsiella
36	27	26.7	26	11	Q61675	Q61675 mus musculu
37	27	26.7	27	12	Q37245	Q37245 hepatitis c
38	27	26.7	28	2	Q9F7C1	Q9F7C1 serratia ma
39	27	26.7	29	4	Q9UCK4	Q9UCK4 homo sapien
40	27	26.7	30	2	Q9L8F9	Q9L8F9 klebsiella
41	27	26.7	30	2	Q9L8I9	Q9L8I9 enterobacte
42	26.5	26.2	28	11	Q61971	Q61971 mus musculu
43	26	25.7	12	4	Q9BZ49	Q9BZ49 homo sapien
44	26	25.7	21	2	Q9S1C0	Q9S1C0 porphyromon
45	26	25.7	22	12	Q69350	Q69350 herpes simp

ALIGNMENTS

RESULT 1

Q16167 PRELIMINARY; PRT; 18 AA.

AC Q16167; (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Serum albumin (Fragment).

GN SERUM ALBUMIN.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94294404; PubMed=8022807;

RA Madison J., Galliano M., Watkins S., Minchiotti L., Porta F.,

RA Rossi A., Putnam F.W.;

RT "Genetic variants of human serum albumin in Italy: point mutants and a

RT carboxyl-terminal variant.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:6476-6480(1994).

DR EMBL: S70799; AAB31177.1; -

FT NON_TER 1

SQ SEQUENCE 18 AA; 2089 MW; 2CCF6C5B0B08690B CRC64;

Query Match

Best Local Similarity 33.7%; Score 34; DB 4; Length 18;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 PRKMKGLFSQAKI 17

Db 5 PRRVKNLLQVKL 17

RESULT 2

Q8C837 PRELIMINARY; PRT; 29 AA.

ID Q8C837

AC Q8C837;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SLM-1 (Fragment).
GN KHDRBS2 OR 6330586C16RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR ENBL; AK048524; BAC33360.1; -.
DR MGD; MGI:2159649; Khdrbs2.
FT NON_TER 1
FT SEQUENCE 29 AA; 3348 MW; 2808CC988D424925 CRC64;
SQ
Query Match 32.7%; Score 33; DB 11; Length 29;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 PPRKMKG 10
Db ||||| :
12 PPRRSARG 19
RESULT 3
QY 092027 PRELIMINARY; PRT; 16 AA.
ID Q9TQZ7
AC Q9TQZ7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Factor H (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE
RX MEDLINE=96202005; PubMed=8615824;
RA Soames C.J., Day A.J., Sim R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved
RT in the interaction with complement component C3b."
RL Biochem. J. 315:523-531 (1996).
SQ SEQUENCE 16 AA; 1871 MW; 7052751E485EC825 CRC64;
Query Match 31.7%; Score 32; DB 6; Length 16;
Best Local Similarity 54.5%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 3 PPRKMKGLFS 13
Db ||||| :
6 PPRKETEILS 16
RESULT 4
QY 037240 PRELIMINARY; PRT; 27 AA.
ID Q37240
AC Q37240
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein 2 (Genome polyprotein) (Fragment).
GN E2.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11;
RX MEDLINE=98105815; PubMed=9445070;
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
RT "Genetic diversity and tissue compartmentalization of the hepatitis C
RT virus genome in blood mononuclear cells, liver, and serum from chronic
RT hepatitis C patients."
RL J. Virol. 72:1640-1646 (1998).
DR ENBL; AF018349; AAC03636.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
FT SEQUENCE 27 AA; 2826 MW; 4C7A918237CB80C2 CRC64;
SQ
Query Match 31.7%; Score 32; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 4 PPRKMKGLFS 13
Db ||||| :
12 PTQKLVGLFS 21
RESULT 5
QY 037239 PRELIMINARY; PRT; 27 AA.
ID Q37239
AC Q37239
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein 2 (Genome polyprotein) (Fragment).
GN E2.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=10;
RX MEDLINE=98105815; PubMed=9445070;
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
RT "Genetic diversity and tissue compartmentalization of the hepatitis C
RT virus genome in blood mononuclear cells, liver, and serum from chronic
RT hepatitis C patients."
RL J. Virol. 72:1640-1646 (1998).
DR ENBL; AF018348; AAC03635.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; 1.
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
FT SEQUENCE 27 AA; 2757 MW; 4C7A9183ECBB80C2 CRC64;
SQ
Query Match 31.7%; Score 32; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 4 PPRKMKGLFS 13

```
Db          |.:|:|||||
            12 PTQKLVGLFS 21

RESULT 6
O37244      PRELIMINARY;      PRT;      27 AA.
AC O37244;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein 2 (Genome polyprotein) (Fragment).
GN E2.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4;
RX MEDLINE=98105815; PubMed=9445070;
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
RT "Genetic diversity and tissue compartmentalization of the hepatitis C
RT virus genome in blood mononuclear cells, liver, and serum from chronic
RT hepatitis C patients.";
RL J. Virol. 72:1640-1646(1998).
DR EMBL; AF018353; AAC03640.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; I.
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2757 MW; 4C7A9183ECBB80C2 CRC64;

Query Match          31.7%; Score 32; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY          4 PPRKMKGLFS 13
            |.:|:|||||
            12 PTQKLVGLFS 21

Db          |.:|:|||||
            12 PTQKLVGLFS 21

RESULT 7
O37242      PRELIMINARY;      PRT;      27 AA.
AC O37242;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein 2 (Genome polyprotein) (Fragment).
GN E2.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=15;
RX MEDLINE=98105815; PubMed=9445070;
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
RT "Genetic diversity and tissue compartmentalization of the hepatitis C
RT virus genome in blood mononuclear cells, liver, and serum from chronic
RT hepatitis C patients.";
RL J. Virol. 72:1640-1646(1998).
DR EMBL; AF018351; AAC03638.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.

Query Match          31.7%; Score 32; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY          4 PPRKMKGLFS 13
            |.:|:|||||
            12 PTQKLVGLFS 21

Db          |.:|:|||||
            12 PTQKLVGLFS 21

RESULT 8
O37222      PRELIMINARY;      PRT;      27 AA.
AC O37222;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein 2 (Genome polyprotein) (Fragment).
GN E2.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A;
RX MEDLINE=98105815; PubMed=9445070;
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
RT "Genetic diversity and tissue compartmentalization of the hepatitis C
RT virus genome in blood mononuclear cells, liver, and serum from chronic
RT hepatitis C patients.";
RL J. Virol. 72:1640-1646(1998).
DR EMBL; AF018331; AAC03611.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR002531; HCV NS1.
DR Pfam; PF01560; HCV NS1; I.
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2757 MW; 4C7A9183ECBB80C2 CRC64;

Query Match          31.7%; Score 32; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY          4 PPRKMKGLFS 13
            |.:|:|||||
            12 PTQKLVGLFS 21

Db          |.:|:|||||
            12 PTQKLVGLFS 21

RESULT 9
O37241      PRELIMINARY;      PRT;      27 AA.
AC O37241;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein 2 (Genome polyprotein) (Fragment).
GN E2.
OS Hepatitis C virus.
```

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13;
 RX MEDLINE=98105815; PubMed=9445070;
 RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
 RT "Genetic diversity and tissue compartmentalization of the hepatitis C
 virus genome in blood mononuclear cells, liver, and serum from chronic
 hepatitis C patients.";
 RL J. Virol. 72:1640-1646(1998).
 DR EMBL; AF018350; AAC03637.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR002531; HCV NS1.
 DR Pfam; PF01560; HCV NS1; I.
 DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2757 MW; 4C7A9183ECBB80C2 CRC64;

Query Match 31.7%; Score 32; DB 12; Length 27;
 Best Local Similarity 60.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPRKMKGLFS 13
 | : : ||||
 Db 12 PTQKLVGLFS 21

RESULT 10

O37238 PRELIMINARY; PRT; 27 AA.
 AC O37238;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope protein 2 (Genome polyprotein) (Fragment).
 GN E2.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9;
 RX MEDLINE=98105815; PubMed=9445070;
 RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
 RT "Genetic diversity and tissue compartmentalization of the hepatitis C
 virus genome in blood mononuclear cells, liver, and serum from chronic
 hepatitis C patients.";
 RL J. Virol. 72:1640-1646(1998).
 DR EMBL; AF018347; AAC03626.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR002531; HCV NS1.
 DR Pfam; PF01560; HCV NS1; I.
 DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2757 MW; 4C7A9183ECBB80C2 CRC64;

Query Match 31.7%; Score 32; DB 12; Length 27;
 Best Local Similarity 60.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPRKMKGLFS 13
 | : : ||||
 Db 12 PTQKLVGLFS 21

RESULT 11

O37237 PRELIMINARY; PRT; 27 AA.
 AC O37237;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope protein 2 (Genome polyprotein) (Fragment).
 GN E2.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8;
 RX MEDLINE=98105815; PubMed=9445070;
 RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
 RT "Genetic diversity and tissue compartmentalization of the hepatitis C
 virus genome in blood mononuclear cells, liver, and serum from chronic
 hepatitis C patients.";
 RL J. Virol. 72:1640-1646(1998).
 DR EMBL; AF018346; AAC03625.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR002531; HCV NS1.
 DR Pfam; PF01560; HCV NS1; I.
 DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KW Polyprotein; Transmembrane.
 FT NON_TER 1
 FT NON_TER 27
 FT NON_TER 27
 SQ SEQUENCE 27 AA; 2757 MW; 4C7A9183ECBB80C2 CRC64;

Query Match 31.7%; Score 32; DB 12; Length 27;
 Best Local Similarity 60.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPRKMKGLFS 13
 | : : ||||
 Db 12 PTQKLVGLFS 21

RESULT 12

O37243 PRELIMINARY; PRT; 27 AA.
 AC O37243;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope protein 2 (Genome polyprotein) (Fragment).
 GN E2.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17;
 RX MEDLINE=98105815; PubMed=9445070;
 RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
 RT "Genetic diversity and tissue compartmentalization of the hepatitis C
 virus genome in blood mononuclear cells, liver, and serum from chronic
 hepatitis C patients.";
 RL J. Virol. 72:1640-1646(1998).
 DR EMBL; AF018352; AAC03639.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002531; HCV NS1.
DR Pfam: PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
FT NON_TER 27 27
FT NON_TER 1 1
SQ SEQUENCE 27 AA; 2757 MW; 4C7A9183ECBB80C2 CRC64;
Query Match 31.7%; Score 32; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 4 PPRKMKGLFS 13
Db 12 PTQKLVGLFS 21
RESULT 13
C37223 PRELIMINARY; PRT; 27 AA.
AC O37223;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope protein 2 (Genome polyprotein) (Fragment).
GN E2.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX MEDLINE=98105815; PubMed=9445070;
RA Navas S., Martin J., Quiroga J.A., Castillo I., Carreno V.;
RT "Genetic diversity and tissue compartmentalization of the hepatitis C virus genome in blood mononuclear cells, liver, and serum from chronic hepatitis C patients.";
RT J. Virol. 72:1640-1646(1998).
DR EMBL: AF018332; AAC03612.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR002531; HCV NS1.
DR Pfam: PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW polyprotein; Transmembrane.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2757 MW; 4C7A9183ECBB80C2 CRC64;
Query Match 31.7%; Score 32; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 4 PPRKMKGLFS 13
Db 12 PTQKLVGLFS 21
RESULT 14
Q8TE34 PRELIMINARY; PRT; 24 AA.
AC Q8TE34;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Nuclear autoantigen (Fragment).
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91079525; PubMed=2258622;
RA Szosteck C., Guldner H.H., Netter H.J., Will H.;
RT "Isolation and characterization of cDNA encoding a human nuclear antigen predominantly recognized by autoantibodies from patients with primary biliary cirrhosis.";
RT J. Immunol. 145:4338-4347(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98075076;
RA Sternsdorf T., Jensen K., Will H.;
RT "Evidence for covalent modification of the nuclear dot-associated proteins PML and Sp100 by PIC1/SUMO-1.";
RT J. Cell Biol. 139:1621-1634(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Guldner H.H., Szosteck C., Will H.;
RT "Alternative splice variants of Sp100.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: I79986; AAL7741.1; -.
FT NON_TER 1 1
SQ SEQUENCE 24 AA; 2856 MW; 801ECB4337635ED8 CRC64;
Query Match 30.7%; Score 31; DB 4; Length 24;
Best Local Similarity 38.5%; Pred. No. 5e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 3 PPRKMKGLFSQA 15
Db 12 PQPQPGLLRQS 24
RESULT 15
Q9BZF8 PRELIMINARY; PRT; 30 AA.
AC Q9BZF8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE NF-E2-related factor 2 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kneé R., Nakamura K., Michalak M.;
RT "Overexpression of calreticulin induces Nrf2 mRNA in HEK 293 cells.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF323119; AAK01657.1; -.
FT NON_TER 30 30
SQ SEQUENCE 30 AA; 3509 MW; 562F233C06A167FA CRC64;
Query Match 30.7%; Score 31; DB 4; Length 30;
Best Local Similarity 50.0%; Pred. No. 6.2e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 4; Gaps 1;
QY 2 LPPPRKMKGLFSQAKISL 19
Db 6 LPPPP---GLPSPQDDMDL 19
Search completed: July 15, 2004, 14:50:55
Job time : 27.7333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: July 15, 2004, 14:47:02 ; Search time 12.4 Seconds
(without alignments)
83.268 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101
Sequence: 1 VLPFRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 204689

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	44.6	23	1	US-07-800-364B-4
2	45	44.6	23	5	PCT-US91-03388-4
3	37	36.6	27	4	US-09-205-258-856
4	33	32.7	16	3	US-08-602-999A-199
5	33	32.7	16	4	US-09-500-124-199
6	33	32.7	27	1	US-08-262-037-107
7	32	31.7	8	3	US-08-444-818-589
8	32	31.7	9	2	US-08-146-028-452
9	32	31.7	9	2	US-08-146-028-453
10	32	31.7	9	3	US-08-723-425A-452
11	32	31.7	9	3	US-08-723-425A-453
12	32	31.7	9	3	US-09-112-206-452
13	32	31.7	9	3	US-09-112-206-453
14	32	31.7	9	4	US-09-790-497A-400
15	32	31.7	9	4	US-09-790-497A-401
16	32	31.7	9	4	US-09-790-497A-576
17	32	31.7	9	4	US-09-576-824A-400
18	32	31.7	9	4	US-09-576-824A-401
19	32	31.7	11	4	US-09-576-824A-576
20	32	31.7	13	2	US-08-572-951-13
21	32	31.7	20	2	US-08-466-975A-20
22	32	31.7	20	2	US-08-391-671A-20
23	32	31.7	20	3	US-08-467-902A-20
24	32	31.7	20	3	US-08-275-265-20
25	32	31.7	20	4	US-08-850-328-13
26	32	31.7	20	4	US-09-941-611-20
27	32	31.7	20	4	US-09-790-497A-63

28	32	31.7	20	4	US-09-790-497A-103	Sequence 103, App
29	32	31.7	22	2	US-08-146-028-63	Sequence 63, Appl
30	32	31.7	22	2	US-08-146-028-103	Sequence 103, App
31	32	31.7	22	3	US-08-723-425A-63	Sequence 63, Appl
32	32	31.7	22	3	US-08-723-425A-103	Sequence 103, App
33	32	31.7	22	3	US-09-112-206-63	Sequence 63, Appl
34	32	31.7	22	3	US-09-112-206-103	Sequence 103, App
35	32	31.7	22	4	US-09-576-824A-63	Sequence 63, Appl
36	32	31.7	22	4	US-09-576-824A-103	Sequence 103, App
37	30	29.7	11	2	US-07-737-371E-19	Sequence 19, Appl
38	30	29.7	16	1	US-07-942-245-210	Sequence 210, App
39	30	29.7	16	3	US-08-602-999A-344	Sequence 344, App
40	30	29.7	16	4	US-09-500-124-344	Sequence 344, App
41	30	29.7	16	6	5378805-5	Patent No. 5378805
42	30	29.7	20	4	US-09-439-313-567	Sequence 567, App
43	30	29.7	21	4	US-09-566-876-12	Sequence 12, Appl
44	30	29.7	22	1	US-08-468-514-11	Sequence 11, Appl
45	30	29.7	29	4	US-09-566-876-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-07-800-364B-4
; Sequence 4, Application US/07800364B
; Patent No. 5688678
; GENERAL INFORMATION:
; APPLICANT: Hewick, Rodney M.
; APPLICANT: Wang, Jack H.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/800,364B
; FILING DATE: 26-NOV-1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI 5182A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Bos taurus
; TISSUE TYPE: Bone
; US-07-800-364B-4

Query Match 44.6%; Score 45; DB 1; Length 23;
Best Local Similarity 63.6%; Pred. No. 0.51;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY      2 LPPPRKMKGLF 12
Db      4 LPPPNKLPGLF 14

RESULT 2
PCT US91-03388-4
; Sequence 4, Application PC/TUS9103388
; GENERAL INFORMATION:
; APPLICANT: Hewick, Rodney M.
; APPLICANT: Wang, Jack H.
; TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/03388
; FILING DATE: 19910515
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GIS182X-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Bos taurus
; TISSUE TYPE: Bone
PCT US91-03388-4

Query Match 44.6%; Score 45; DB 5; Length 23;
Best Local Similarity 63.6%; Pred. No. 0.51;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 LPPPRKMKGLF 12
Db      4 LPPPNKLPGLF 14

RESULT 3
US 09-205-258-856
; Sequence 856, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921

```

```

; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 856
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-856

```

Query Match 36.6%; Score 37; DB 4; Length 27;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 6; Conservative 3; Mismatches 3; Indels

```

QY      2 LPPPRKMKGLFS 13
         :||| ||| :
Db      16 VPPPLKMPGVLA 27

```

```

RESULT 4
US-08-602-999A-199
; Sequence 199, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467

```

```

Query Match      32.7%; Score 33; DB 3; Length 16;
Best Local Similarity 75.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 2; Indels

Qy      3 PPRKMKG 10
      |||||

```

Db 6 PP1PMKG 13

RESULT 5

US-09-500-124-199

; Sequence 199, Application US/09500124

; Patent No. 6432920

; GENERAL INFORMATION:

; APPLICANT: SPARKS, Andrew B.

; APPLICANT: KAY, Brian K.

; APPLICANT: THORN, Judith M.

; APPLICANT: QUILLIAM, Lawrence A.

; APPLICANT: DER, Channing J.

; APPLICANT: FOLKES, Dana M.

; APPLICANT: RIDER, James E.

; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF

; TITLE OF INVENTION: ISOLATING AND USING SAME

; NUMBER OF SEQUENCES: 467

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/500,124

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/602,999

; FILING DATE: 16-FEB-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 1101-202

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 199:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 16 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

US-09-500-124-199

Query Match	32.7%	Score 33;	DB 4;	Length 16;
Best Local Similarity	75.0%	Pred. No. 36;		
Matches 6;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps 0;				

Qy 3 P P P R K M K G 10
Db 6 P P P I P M K G 13

Db 6 PPPIPMXG 13

RESULT 6
US-08-262-037-107
; Sequence 107, Application US/08262037
; Patent No. 5747239
; GENERAL INFORMATION:
; APPLICANT: Chang Yi Wang and Barbara Hosein
; TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
; TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
; TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
; TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVE.
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/719,819
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/667,275
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 07-Feb-1991
APPLICATION NUMBER: 07/558,799
FILING DATE: 26-July-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-April-1990
ATTORNEY/AGENT INFORMATION:
NAME: Maria C. H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4043 US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: Amino acid
STRANDEDNESS:
TOPOLOGY: Unknown
US-08-262-037-107

Query Match 32.7%; Score 33; DB 1; Length 27;
Best Local Similarity 62.5%; Pred. No. 64;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LPPPRKMK 9
Db 19 VPPPRKKR 26

RESULT 7
US-08-444-818-589
Sequence 589, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David V.
APPLICANT: Ruter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 589:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-444-818-589

Query Match 31.7%; Score 32; DB 3; Length 8;
Best Local Similarity 71.4%; Pred. No. 3e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PPRKMK 9
Db 1 PPRKKR 7

RESULT 8
US-08-146-028-452
Sequence 452, Application US/08146028
Patent No. 5891640
GENERAL INFORMATION:
APPLICANT:

TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/146,028
INFORMATION FOR SEQ ID NO: 452:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-146-028-452

Query Match 31.7%; Score 32; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 3e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPPRK 7
Db 4 VPPPRK 9

RESULT 9
US-08-146-028-453
Sequence 453, Application US/08146028
Patent No. 5891640
GENERAL INFORMATION:

;; APPLICANT: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA: US/08/146,028
;; APPLICATION NUMBER: US/08/146,028
;; INFORMATION FOR SEQ ID NO: 453:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-146-028-453

Query Match 31.7%; Score 32; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPPPRK 7
Db 3 VPPPRK 8

RESULT 10

US-08-723-425A-452
; Sequence 452, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:

;; APPLICANT: DELEYS, ROBERT
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
;; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
;; NUMBER OF SEQUENCES: 453

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: NIXON & VANDERHYE, P.C.
;; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
;; CITY: Arlington
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22201

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/723,425A
;; FILING DATE:

;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SADOFF, B.J.
;; REGISTRATION NUMBER: 36,663
;; REFERENCE/DOCKET NUMBER: 1487-13
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-816-4000
;; TELEFAX: 703-816-4100

;; INFORMATION FOR SEQ ID NO: 452:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide

US-08-723-425A-452

Query Match 31.7%; Score 32; DB 3; Length 9;
Best Local Similarity 83.3%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPPPRK 7
Db 4 VPPPRK 9

RESULT 11

US-08-723-425A-453
; Sequence 453, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:

;; APPLICANT: DELEYS, ROBERT
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
;; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
;; NUMBER OF SEQUENCES: 453

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: NIXON & VANDERHYE, P.C.
;; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
;; CITY: Arlington
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22201

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/723,425A
;; FILING DATE:

;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SADOFF, B.J.
;; REGISTRATION NUMBER: 36,663
;; REFERENCE/DOCKET NUMBER: 1487-13
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-816-4000
;; TELEFAX: 703-816-4100

;; INFORMATION FOR SEQ ID NO: 453:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-723-425A-453

Query Match 31.7%; Score 32; DB 3; Length 9;
Best Local Similarity 83.3%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LPPPRK 7
Db 3 VPPPRK 8

RESULT 12

US-08-112-206-452
; Sequence 452, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:

;; APPLICANT: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES, ...

1 TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
2 NUMBER OF SEQUENCES: 453
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: Floppy disk
5 COMPUTER: IBM PC compatible
6 OPERATING SYSTEM: PC-DOS/MS-DOS
7 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
8 CURRENT APPLICATION DATA:
9 APPLICATION NUMBER: US/09/112,206
10 FILING DATE:
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 08/146,028
13 FILING DATE:
14 INFORMATION FOR SEQ ID NO: 452:
15 SEQUENCE CHARACTERISTICS:
16 LENGTH: 9 amino acids
17 TYPE: amino acid
18 STRANDEDNESS: single
19 TOPOLOGY: linear
20 MOLECULE TYPE: peptide
21 US-09-112-206-452

Query Match 31.7%; Score 32; DB 3; Length 9;
Best Local Similarity 83.3%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPPRK 7
:|||||
Db 4 VPPPRK 9

RESULT 13
US-09-112-206-453
1 Sequence 453, Application US/09112206
2 Patent No. 6210903
3 GENERAL INFORMATION:
4 APPLICANT:
5 TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
6 CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
7 TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
8 TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
9 TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
10 NUMBER OF SEQUENCES: 453
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/09/112,206
18 FILING DATE:
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: US 08/146,028
21 FILING DATE:
22 INFORMATION FOR SEQ ID NO: 453:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 9 amino acids
25 TYPE: amino acid
26 STRANDEDNESS: single
27 TOPOLOGY: linear
28 MOLECULE TYPE: peptide
29 US-09-112-206-453

Query Match 31.7%; Score 32; DB 3; Length 9;
Best Local Similarity 83.3%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPPRK 7
:|||||
Db 3 VPPPRK 8

RESULT 14

US-09-790-497A-400
1 Sequence 400, Application US/09790497A
2 Patent No. 6649735
3 GENERAL INFORMATION:
4 APPLICANT: De Leys, Robert
5 TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
6 TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
7 TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
8 TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
9 TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
10 FILE REFERENCE: 2752-16
11 CURRENT APPLICATION NUMBER: US/09/790,497A
12 CURRENT FILING DATE: 2001-02-23
13 PRIOR APPLICATION NUMBER: 09/576,824
14 PRIOR FILING DATE: 2000-05-23
15 PRIOR APPLICATION NUMBER: 08/723,425
16 PRIOR FILING DATE: 1996-09-30
17 PRIOR APPLICATION NUMBER: 09/146,028
18 PRIOR FILING DATE: 1993-11-22
19 PRIOR APPLICATION NUMBER: PCT/EP93/00517
20 PRIOR FILING DATE: 1993-03-08
21 PRIOR APPLICATION NUMBER: EP 92400598.6
22 PRIOR FILING DATE: 1992-03-06
23 NUMBER OF SEQ ID NOS: 600
24 SOFTWARE: PatentIn Ver. 2.1
25 SEQ ID NO 400
26 LENGTH: 9
27 TYPE: PRT
28 ORGANISM: Hepatitis C virus
29 US-09-790-497A-400

Query Match 31.7%; Score 32; DB 4; Length 9;
Best Local Similarity 83.3%; Pred. No. 3e+05; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LPPPRK 7
:|||||
Db 4 VPPPRK 9

RESULT 15

US-09-790-497A-401
1 Sequence 401, Application US/09790497A
2 Patent No. 6649735
3 GENERAL INFORMATION:
4 APPLICANT: De Leys, Robert
5 TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
6 TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
7 TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
8 TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
9 TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
10 FILE REFERENCE: 2752-16
11 CURRENT APPLICATION NUMBER: US/09/790,497A
12 CURRENT FILING DATE: 2001-02-23
13 PRIOR APPLICATION NUMBER: 09/576,824
14 PRIOR FILING DATE: 2000-05-23
15 PRIOR APPLICATION NUMBER: 08/723,425
16 PRIOR FILING DATE: 1996-09-30
17 PRIOR APPLICATION NUMBER: 09/146,028
18 PRIOR FILING DATE: 1993-11-22
19 PRIOR APPLICATION NUMBER: PCT/EP93/00517
20 PRIOR FILING DATE: 1993-03-08
21 PRIOR APPLICATION NUMBER: EP 92400598.6
22 PRIOR FILING DATE: 1992-03-06
23 NUMBER OF SEQ ID NOS: 600
24 SOFTWARE: PatentIn Ver. 2.1
25 SEQ ID NO 401
26 LENGTH: 9
27 TYPE: PRT
28 ORGANISM: Hepatitis C virus
29 US-09-790-497A-401

Query Match 31.7%; Score 32; DB 4; Length 9;
 Best Local Similarity 83.3%; Pred. No. 3e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LPPPRK 7
 Db 3 VPPPRK 8

Search completed: July 15, 2004, 15:13:19
 Job time : 14.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:11:48 ; Search time 32.5333 Seconds
(without alignments)
192.148 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101

Sequence: 1 VLPPPRKMGFLFSQAKISLP 20

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 288454

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	10	US-09-171-432A-42
2	65	64.4	25	10	US-09-171-432A-47
3	42	41.6	20	10	US-09-171-432A-43
4	39	38.6	20	10	US-09-171-432A-41
5	38	37.6	28	12	US-10-424-599-236962
6	37	36.6	25	14	US-10-029-386-34051
7	37	36.6	27	10	US-09-933-767-856
8	37	36.6	27	12	US-10-004-860-856
9	37	36.6	27	14	US-10-023-282-856
10	36	35.6	29	9	US-09-872-854-17
11	35	34.7	29	9	US-09-872-854-18
12	33	32.7	16	14	US-10-161-791-199
13	33	32.7	25	16	US-10-109-048-909
14	33	32.7	30	9	US-09-864-761-46386
15	33	32.7	30	12	US-10-296-734-716

Sequence 718, App
Sequence 4, Appli
Sequence 20, Appl
Sequence 20, Appl
Sequence 43490, A
Sequence 37832, A
Sequence 323, App
Sequence 325, App
Sequence 2, Appli
Sequence 344, App
Sequence 567, App
Sequence 12, Appl
Sequence 3, Appli
Sequence 92, Appl
Sequence 305, App
Sequence 44248, A
Sequence 566, App
Sequence 566, App
Sequence 47733, A
Sequence 38476, A
Sequence 13, Appl
Sequence 714, App
Sequence 302, App
Sequence 302, App
Sequence 15, Appl
Sequence 563, App
Sequence 564, App
Sequence 565, App
Sequence 8, Appli
Sequence 8, Appli

16 33 32.7 30 12 US-10-296-734-718
17 32 31.7 17 15 US-10-149-536-4
18 32 31.7 20 9 US-09-941-611-20
19 32 31.7 20 14 US-10-044-995-20
20 32 31.7 22 9 US-09-864-761-43490
21 32 31.7 23 9 US-09-864-761-37832
22 31 30.7 26 14 US-10-174-410-323
23 31 30.7 26 14 US-10-174-410-325
24 31 30.7 29 16 US-10-343-663A-2
25 30 29.7 16 14 US-10-161-791-344
26 30 29.7 20 14 US-10-010-940-567
27 30 29.7 21 13 US-10-082-659-12
28 30 29.7 21 14 US-10-097-175-3
29 30 29.7 21 14 US-10-097-175-92
30 30 29.7 23 14 US-10-174-410-305
31 30 29.7 25 9 US-09-864-761-44248
32 30 29.7 25 12 US-09-925-298-566
33 30 29.7 25 14 US-10-102-806-566
34 30 29.7 27 9 US-09-864-761-47733
35 30 29.7 29 9 US-09-864-761-38476
36 30 29.7 29 13 US-10-082-659-13
37 30 29.7 30 12 US-10-296-734-714
38 29.5 29.2 28 12 US-10-653-595-302
39 29.5 29.2 28 12 US-09-397-945-302
40 29 28.7 12 16 US-10-688-100-15
41 29 28.7 15 16 US-10-416-249-563
42 29 28.7 15 16 US-10-416-249-564
43 29 28.7 15 16 US-10-416-249-565
44 29 28.7 16 9 US-09-813-383-8
45 29 28.7 16 9 US-09-813-463A-8

ALIGNMENTS

RESULT 1

; Sequence 42, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Khudiyakov, Yuri A.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171.432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label= YK-1318
US-09-171-432A-42

Query Match 100.0%; Score 101; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 1 VLPPPRKMKGLFSQAKISLF 20

RESULT 2

US-09-171-432A-47
; Sequence 47, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,457
REFERENCE/DOCKET NUMBER: 03063-0231US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499

INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:

NAME/KEY: Peptide
LOCATION: 1..25
OTHER INFORMATION: /label= YK-1665
US-09-171-432A-47

Query Match 64.4%; Score 65; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLF 12
DB 14 VLPPPRKMKGLF 25

RESULT 3

US-09-171-432A-43
; Sequence 43, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:

; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499

INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:

NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label= YK-1319
US-09-171-432A-43

Query Match 41.6%; Score 42; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FSQAKISLF 20
DB 1 FSQAKISLF 9

RESULT 4

US-09-171-432A-41
; Sequence 41, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:

; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88

```
/
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Kilpatrick Stockton LLP
/ STREET: 3424 Peachtree Road, N.E.
/ CITY: Atlanta
/ STATE: Georgia
/ COUNTRY: USA
/ ZIP: 30326
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/171,432A
/ FILING DATE: 23-NOV-1998
/
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Greene, Jamie L.
/ REGISTRATION NUMBER: 32,467
/ REFERENCE/DOCKET NUMBER: 03063-0231US
/
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (404) 949-2400
/ TELEFAX: (404) 949-2499
/
/ INFORMATION FOR SEQ ID NO: 41:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/
/ FEATURE:
/ NAME/KEY: Peptide
/ LOCATION: 1..20
/ OTHER INFORMATION: /label= YK-1317
/
US-09-171-432A-41

Query Match 38.6%; Score 39; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRK 7
Db 14 VLPPRK 20

RESULT 5
US-10-424-599-236962
/ Sequence 236962, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 236962
/ LENGTH: 28
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MPT3847_56001C.1.pap
US-10-424-599-236962

Query Match 37.6%; Score 38; DB 12; Length 28;
Best Local Similarity 44.4%; Pred. No. 1e+02;
```

```
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 LPPPRKMKGLFSQAKISL 19
   |||:| |||: |||
Db 3 LPPQKKKNGLLFKPPLSL 20
   |||:| |||: |||

RESULT 6
US-10-029-386-34051
/ Sequence 34051, Application US/10029386
/ Publication No. US20030194704A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
/ FILE REFERENCE: AEOMICA-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 34051
/ LENGTH: 25
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC023120.3
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.44
US-10-029-386-34051

Query Match 36.6%; Score 37; DB 14; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 KMKGLFSQAKISLF 20
   |:|:| |||
Db 10 KLEGITAEKIRLF 23
   |:|:| |||

RESULT 7
US-09-933-767-856
/ Sequence 856, Application US/09933767
/ Publication No. US20030181692A1
/ GENERAL INFORMATION:
/ APPLICANT: Ni et al.
/ TITLE OF INVENTION: 207 Human Secreted Proteins
/ FILE REFERENCE: P2007P2
/ CURRENT APPLICATION NUMBER: US/09/933,767
/ CURRENT FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: PCT/US01/05614
/ PRIOR FILING DATE: 2001-02-21
/ PRIOR APPLICATION NUMBER: 60/184,836
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/193,170
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: 09/205,258
/ PRIOR FILING DATE: 1998-12-04
/ PRIOR APPLICATION NUMBER: PCT/US98/11422
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/048,885
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/049,375
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/048,881
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/048,880
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/048,896
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/049,020
/ PRIOR FILING DATE: 1997-06-06
/ PRIOR APPLICATION NUMBER: 60/048,876
```

;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,895
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,884
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,894
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,971
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,964
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,882
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,899
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,893
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,900
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,901
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,892
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,915
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/049,019
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,970
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,972
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,916
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/049,373
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,875
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/049,374
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,917
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,949
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,974
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,883
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,897
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,898
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,962
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,963
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,877
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,878
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/068,054
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,064
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,053
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/070,923
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/073,160
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: 60/073,159
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: 60/073,165
;; PRIOR FILING DATE: 1998-01-30

;; PRIOR APPLICATION NUMBER: 60/073,164
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: 60/085,925
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/085,921
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/085,923
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/085,922
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/092,921
;; PRIOR FILING DATE: 1998-07-15
;; PRIOR APPLICATION NUMBER: 60/094,657
;; PRIOR FILING DATE: 1998-07-30
;; NUMBER OF SEQ ID NOS: 1245
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 856
;; LENGTH: 27
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-933-767-856

Query Match 36.6%; Score 37; DB 10; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPPRKMKGLFS 13
Db 16 VPPPLKMPGVLA 27

RESULT 8
US-10-004-860-856
;; Sequence 856, Application US/10004860
;; Publication No. US20030085160A1
;; GENERAL INFORMATION:
;; APPLICANT: Young et al.
;; TITLE OF INVENTION: 207 Human Secreted Proteins
;; FILE REFERENCE: P2007P1
;; CURRENT APPLICATION NUMBER: US/10/004,860
;; CURRENT FILING DATE: 2001-12-07
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 1227
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 856
;; LENGTH: 27
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-004-860-856

Query Match 36.6%; Score 37; DB 12; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPPRKMKGLFS 13
Db 16 VPPPLKMPGVLA 27

RESULT 9
US-10-023-282-856
;; Sequence 856, Application US/10023282
;; Publication No. US20030092893A1
;; GENERAL INFORMATION:
;; APPLICANT: Young et al.
;; TITLE OF INVENTION: 207 Human Secreted Proteins
;; FILE REFERENCE: P2007P1
;; CURRENT APPLICATION NUMBER: US/10/023,282
;; CURRENT FILING DATE: 2001-12-20
;; EARLIER APPLICATION NUMBER: 09/205,258
;; EARLIER FILING DATE: 1998-12-04
;; EARLIER APPLICATION NUMBER: PCT/US98/11422
;; EARLIER FILING DATE: 1998-06-04

; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923

; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 856
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-282-856

Query Match 36.6%; Score 37; DB 14; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPPRKMKGLFS 13

DB 16 VPPPLKMPGVLA 27

RESULT 10

US-09-872-864-17

; Sequence 17, Application US/09872864

; Patent No. US2002011305A1

; GENERAL INFORMATION:

; APPLICANT: AMERONGEN, ARIE VAN NIEUW

; APPLICANT: VEERMAN, ENGELMUNDUS CORNELIS IGNATIUS

; APPLICANT: VAN 'T HOF, WILLEM

; APPLICANT: NIBBERING, PETER HENDRICUS

; TITLE OF INVENTION: ANTIVIRAL PEPTIDES

; FILE REFERENCE: EPP-004.01

; CURRENT APPLICATION NUMBER: US/09/872,864

; CURRENT FILING DATE: 2001-06-01

; PRIOR APPLICATION NUMBER: PCT/NL99/00732

; PRIOR FILING DATE: 1999-12-01

; PRIOR APPLICATION NUMBER: NL1010692

; PRIOR FILING DATE: 1998-12-01

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 17

; LENGTH: 29

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: peptide

; NAME/KEY: MOD.RES

; LOCATION: (29)

; OTHER INFORMATION: AMIDATION

US-09-872-864-17

Query Match

Best Local Similarity 35.6%; Score 36; DB 9; Length 29;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 RKMKGFLFSQAKISL 19

DB 12 RKYKRLPKELKFSL 25

RESULT 11

US-09-872-864-18

; Sequence 18, Application US/09872864

; Patent No. US2002011305A1

; GENERAL INFORMATION:

; APPLICANT: AMERONGEN, ARIE VAN NIEUW

; APPLICANT: VEERMAN, ENGELMUNDUS CORNELIS IGNATIUS

; APPLICANT: VAN 'T HOF, WILLEM

; APPLICANT: NIBBERING, PETER HENDRICUS

; TITLE OF INVENTION: ANTIVIRAL PEPTIDES

; FILE REFERENCE: EPP-004.01

;; CURRENT APPLICATION NUMBER: US/09/872,864
;; CURRENT FILING DATE: 2001-06-01
;; PRIOR APPLICATION NUMBER: PCT/NL99/00732
;; PRIOR FILING DATE: 1999-12-01
;; PRIOR APPLICATION NUMBER: NL1010692
;; PRIOR FILING DATE: 1998-12-01
;; NUMBER OF SEQ ID NOS: 20
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 18
;; LENGTH: 29
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; NAME/KEY: MOD RES
;; LOCATION: (29)
;; OTHER INFORMATION: AMIDATION
US-09-872-864-18

Query Match 34.7%; Score 35; DB 9; Length 29;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 RKMKGFLFSQAKISL 19
||| ||| : |||
Db 12 RYKRLFKKLKLSL 25

RESULT 12
US-10-161-791-199
;; Sequence 199, Application US/10161791
;; Publication No. US2003018683A1
;; GENERAL INFORMATION:
;; APPLICANT: SPARKS, Andrew B.
;; APPLICANT: KAY, Brian K.
;; APPLICANT: THORN, Judith M.
;; APPLICANT: QUILLIAM, Lawrence A.
;; APPLICANT: DER, Channing J.
;; APPLICANT: FOLKES, Dana M.
;; APPLICANT: RIDER, James E.
;; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
;; TITLE OF INVENTION: ISOLATING AND USING SAME
;; NUMBER OF SEQUENCES: 467
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/161,791
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/602,999
;; FILING DATE: 16-FEB-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Misrock, S. Leslie
;; REGISTRATION NUMBER: 18,872
;; REFERENCE/DOCKET NUMBER: 1101-202
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-9741/8864
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 199:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 16 amino acids
;; TYPE: amino acid

;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-10-161-791-199

Query Match 32.7%; Score 33; DB 14; Length 16;
Best Local Similarity 75.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPRKMKKG 10
||| |||
Db 6 PPIPMMKG 13

RESULT 13
US-10-109-048-909
;; Sequence 909, Application US/10109048
;; Publication No. US20040107461A1
;; GENERAL INFORMATION:
;; APPLICANT: COMMURI, PADMA
;; APPLICANT: KEELING, PETER L.
;; APPLICANT: RAMIREZ, NONA
;; APPLICANT: MCKEAN, ANGELA
;; APPLICANT: GAO, ZHONG
;; APPLICANT: GUAN, HANPING
;; TITLE OF INVENTION: GLUCAN CHAIN LENGTH DOMAINS
;; FILE REFERENCE: 2461-76
;; CURRENT APPLICATION NUMBER: US/10/109,048
;; CURRENT FILING DATE: 2003-03-04
;; PRIOR APPLICATION NUMBER: 60/279,720
;; PRIOR FILING DATE: 2001-03-30
;; NUMBER OF SEQ ID NOS: 1154
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 909
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Unknown Organism
;; FEATURE:
;; OTHER INFORMATION: Description of Unknown Organism: Accession No. 15672681
US-10-109-048-909

Query Match 32.7%; Score 33; DB 16; Length 25;
Best Local Similarity 46.7%; Pred. No. 5.2e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 RKMKGFLFSQAKISL 20
||| ||| : |||
Db 3 RKMVLPASSECAFF 17

RESULT 14
US-09-864-761-46386
;; Sequence 46386, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27


```

, PRIOR APPLICATION NUMBER: PCT/US01/006666
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006667
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006664
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006669
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006665
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006668
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006663
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006662
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006661
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/006670
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: US 60/234,687
, PRIOR FILING DATE: 2000-09-21
, PRIOR APPLICATION NUMBER: US 09/608,408
, PRIOR FILING DATE: 2000-06-30
, PRIOR APPLICATION NUMBER: US 09/774,203
, PRIOR FILING DATE: 2001-01-29
, NUMBER OF SEQ ID NOS: 49117
, SOFTWARE: Annonam Sequence Listing Engine
, SEQ ID NO 46386
, LENGTH: 30

```

```
Query Match      32.7%; Score 33; DB 9; Length 30;
Best Local Similarity 31.6%; Pred. No. 6.3e+02;
Matches 6; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
```

Qy 1 VLPPPRMKGLFSQAKISL 19
: : | | : | | | : :
Db 7 LIPPSDPLRLVSOAEVEM 25

```

RESULT 15
US-10-296-734-716
; Sequence 716, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramsnaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 716
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: HepC la segment 155
US-10-296-734-716

```

Query Match 32.7%; Score 33; DB 12; Length 30;
Best Local Similarity 62.5%; Pred. NO. 6.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels

QY 2 LPPPRKMK 9
:|||||:
Db 16 VPPPRKKR 23

Search completed: July 15, 2004, 15:22:54
Job time : 33.5333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:36:31 ; Search time 40.1333 Seconds
(without alignments)
140.804 Million cell updates/sec

Title: US-09-171-432a-43
Perfect score: 101
Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 581562

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	20	2 AAW42926	Aaw42926 Immunogen
2	101	100.0	21	4 AAB69443	Aab69443 Synthetic
3	49.5	49.0	20	2 AAW76539	Aaw76539 Plasmid p
4	42	41.6	20	2 AAW42925	Aaw42925 Immunogen
5	42	41.6	21	4 AAB69442	Aab69442 Synthetic
6	34	33.7	20	6 ABJ38218	Abj38218 Human cyt
7	34	33.7	25	5 ABG62573	Abg62573 Eubacteri
8	34	33.7	28	6 ABR81914	AbR81914 Human int
9	33	32.7	10	4 AAB35436	Aab35436 Nascent p
10	33	32.7	10	7 ADA02802	Ada02802 Nascent p
11	33	32.7	10	7 ADC15517	Adc15517 Aminoacyl
12	33	32.7	10	7 ADD23457	Add23457 Breast ca
13	32	31.7	15	2 AAY30497	Aay30497 Mutant se
14	32	31.7	20	3 AAY67695	Aay67695 Peptide #
15	32	31.7	25	2 AAR31936	Aar31936 Human TC-
16	32	31.7	25	4 AAE05138	Aae05138 Synthetic
17	32	31.7	28	6 ABJ18881	Abj18881 Human Bcl
18	32	31.7	30	2 AAR31935	Aar31935 Human TC-
19	32	31.7	30	4 AAE05137	Aae05137 Bioactive
20	31.5	31.2	19	2 AAR70836	Aar70836 Murine JA
21	31.5	31.2	19	3 AAB35717	Aab35717 Jak2 doma
22	31.5	31.2	19	4 AAE00348	Aae00348 Murine (J
23	31.5	31.2	20	2 AAW52280	Aaw52280 Synthetic
24	31.5	31.2	20	2 AAW52298	Aaw52298 Synthetic
25	31	30.7	15	2 AAY30502	Aay30502 Mutant se

RESULT 1				
AAW42926	ID	AAW42926	standard; peptide; 20 AA.	
XX	AC	AAW42926;		
XX	DT	28-APR-1998	(first entry)	
XX	DE	Immunogenic Hepatitis A virus peptide YK-1319.		
XX	KW	Immunogenic peptide; immunogenic epitope; P2A protein; immune response; antibody.		
XX	OS	Synthetic.		
XX	OS	Hepatitis A virus.		
XX	PN	WO9740147-A1.		
XX	PD	30-OCT-1997.		
XX	PF	18-APR-1997; 97WO-US006891.		
XX	PR	19-APR-1996; 96US-0015644P.		
XX	PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
XX	FI	Fields HA, Khudyakov YE;		
XX	DR	WPI; 1997-535831/49.		
XX	PT	Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune response to HAV in a mammal or to detect the presence of antibodies against HAV in a mammal.		
XX	PS	Claim 18; Page 112; 140pp; English.		
XX	CC	Peptides AAW42922-30 are immunogenic peptides corresponding to immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are substantially similar to a portion of the amino acid sequence of the P2A protein of HAV corresponding to amino acids 792-980. The present peptide is derived from amino acids 834-853, and has a reactivity of 27.1% with acute sera. Compositions containing the peptides can be used to induce an immune response to HAV in a mammal. The peptides can also be used to detect the presence of antibodies against HAV in mammalian serum. The peptides can also be used to make an antibody against HAV by administering the peptide to a mammal		
XX	SQ	Sequence 20 AA;		

ALIGNMENTS

Abp83600 G protein
Aay67692 Peptide #
Abj38230 Human cyt
Abg62504 Eubacteri
Abg62273 Eubacteri
Aaw34460 Mouse MHC
Abj38196 Human cyt
Abj38197 Human cyt
Aab26685 Partial s
Aaw18068 Soluble s
Aaw72584 SSSI maiz
Aay30507 Mutant se
Aab87042 Human TAN
Aau86368 Oestrogen
Aab89740 HIV gp120
Aaw74171 H7 loop p
Aay67693 Peptide #
Aau98307 Peptide s
Aari5224 N-termina
Aab89479 HIV gp120

Query Match 100.0%; Score 101; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20
 |||||
 DB 1 FSOAKISLFYTEEHEIMKFS 20

RESULT 2

AAW69443
 ID AAB69443 standard; peptide; 21 AA.

XX AC AAB69443;

XX DT 20-APR-2001 (first entry)

XX DE Synthetic HAV P2A peptide, SEQ ID NO: 43.

XX KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX OS Hepatitis A virus.

XX OS Synthetic.

XX PN WO200105824-A2.

XX PD 25-JAN-2001.

XX PF 14-JUL-2000; 2000WO-US019267.

XX PR 15-JUL-1999; 99US-0144412P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Fields HA, Khudyakov YE;

XX DR WPI; 2001-112681/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.

XX PS Claim 13; Page 95; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IGM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IGM antibody reactivity

XX SQ Sequence 21 AA;

Query Match 100.0%; Score 101; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20
 |||||
 DB 1 FSOAKISLFYTEEHEIMKFS 20

RESULT 3

AAW76539

ID AAW76539 standard; protein; 20 AA.

XX AC AAW76539;

XX DT 11-DEC-1998 (first entry)

XX DE Plasmid pGEM2-HAV-delta2BC protein fragment.

XX KW Picornavirus; structural protein; P1-2A; P1; P3 region; HAV;
 KW hepatitis A virus; viral replication; empty viral particle.

XX OS Synthetic.

XX PN WO9844122-A1.

XX PD 08-OCT-1998.

XX PF 26-MAR-1998; 98WO-DR000879.

XX PR 27-MAR-1997; 97DE-01012899.

XX PA (NOVE-) NOVEMBER NOVUS MEDICATUS BERTLING GES MO.

XX PI Probst C;

XX DR WPI; 1998-557117/47.

XX DR N-PSDB; AAV61866.

XX Recombinant production of picorna virus - by co-expressing structural
 PT protein precursors P1 and P3, in cis or trans, particularly to produce
 PT hepatitis A particles for use in vaccines.

XX PS Disclosure; Fig 3; 25pp; German.

XX This sequence represents a fragment of the plasmid pGEM2-HAV-delta2BC
 CC which is used in a novel method which results in the recombinant
 CC production of picornavirus particles, their precursors or partial
 CC derivatives and involves co-expressing structural protein precursor
 CC molecules P1-2A or P1, with the complete P3 region (3ABCD), in cis or in
 CC trans. This invention has particular use with hepatitis A virus (HAV)
 CC particles, which does not involve expressing the P2B or 2C
 CC proteins, is more efficient and excludes the possibility of viral
 CC replication. Using the complete P3 region, rather than just 3C and 3D,
 CC significantly improves production of empty viral particles

XX SQ Sequence 20 AA;

Query Match 49.0%; Score 49.5; DB 2; Length 20;
 Best Local Similarity 75.0%; Pred. No. 0.35;
 Matches 12; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 3 QAKISLFYTEEHEIMK 18
 |||||
 DB 1 QANISLFYTEE-EFME 15

RESULT 4

AAW42925

ID AAW42925 standard; peptide; 20 AA.

XX AC AAW42925;

XX DT 28-APR-1998 (first entry)

XX DE Immunogenic Hepatitis A virus peptide YK-1318.

XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
 KW antibody.

XX OS Synthetic.

OS Hepatitis A virus.

XX PN WO9740147-A1.
 XX PD 30-OCT-1997.
 XX PF 18-APR-1997; 97WO-US006891.
 XX PR 19-APR-1996; 96US-0015644P.
 XX FA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Fields HA, Khudyakov YE;
 XX XPI; 1997-535831/49.
 XX DR
 XX PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.
 XX PS Claim 18; Page 112; 140pp; English.
 XX XQ Peptides AAM42922-30 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the p2A
 CC protein of HAV corresponding to amino acids 792-980. The present peptide
 CC is derived from amino acids 823-842, and has a reactivity of 31.3% with
 CC acute sera. Compositions containing the peptides can be used to induce an
 CC immune response to HAV in a mammal. The peptides can also be used to
 CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal
 XX XQ Sequence 20 AA;
 Query Match 41.6%; Score 42; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FSQAKISLF 9
 Db |||||
 12 FSQAKISLF 20
 RESULT 5
 AAB69442
 ID AAB69442 standard; peptide; 21 AA.
 XX AC AAB69442;
 XX DT 20-APR-2001 (first entry)
 XX DE Synthetic HAV P2A peptide, SEQ ID NO: 42.
 XX KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 XX OS Hepatitis A virus.
 XX OS Synthetic.
 XX PN WO200105824-A2.
 XX PD 25-JAN-2001.
 XX PF 14-JUL-2000; 2000WO-US019267.
 XX PR 15-JUL-1999; 99US-0144412P.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Fields HA, Khudyakov YE;
 XX DR WPI; 2001-112681/12.

PT Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 XX PS Claim 13; Page 95; 130pp; English.
 XX XQ The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IgM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IgM antibody reactivity
 XX XQ Sequence 21 AA;
 Query Match 41.6%; Score 42; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 6.4; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FSQAKISLF 9
 Db |||||
 12 FSQAKISLF 20
 RESULT 6
 ABJ38218
 ID ABJ38218 standard; peptide; 20 AA.
 XX AC ABJ38218;
 XX DT 22-MAY-2003 (first entry)
 XX DE Human cytomegalovirus CTL epitope peptide SEQ ID No 276.
 XX KW Virucide; immunostimulant; cytotoxic T-lymphocyte; CTL; epitope; antigen;
 KW human cytomegalovirus; HCMV; pp28; pp50; pp65; pp71; pp150; gB; gH; IE-1;
 KW IE-2; US2; US3; US6; US11; ULI8; cytomegalovirus infection; pregnancy;
 KW transplantation.
 XX OS Human cytomegalovirus.
 XX PN WO2003000720-A1.
 XX PD 03-JAN-2003.
 XX PF 26-JUN-2002; 2002WO-AU000829.
 XX PR 26-JUN-2001; 2001AU-00005931.
 XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
 XX PI Khanna R, Elkington RA, Walker SJ;
 XX XPI; 2003-300379/29.
 XX DR New human cytomegalovirus (CMV) cytotoxic T-cell epitope peptide, useful
 XX for diagnosing, preventing or treating CMV infection, comprises pp28,
 PT pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or ULI8.
 PS Claim 70; Page 124; 308pp; English.
 XX The invention relates to a novel isolated peptide comprising one or more
 CC cytotoxic T-lymphocyte (CTL) epitope(s) of the same or different antigen
 CC of a human cytomegalovirus (HCMV). The HCMV antigens can comprise pp28,

CC pp50, pp65, pp71, pp150, gB, gH, IE-1, IE-2, US2, US3, US6, US11 or UL18.
 CC The peptide comprises a sequence of about 9-20 contiguous amino acids of
 CC the antigen. The peptide epitopes are useful in diagnosing, preventing or
 CC treating cytomegalovirus infection in humans, and in monitoring immune
 CC responses in various clinical settings (e.g. transplantation or
 CC pregnancy). This sequence represents a human cytomegalovirus CTL epitope
 CC peptide of the invention

XX SQ Sequence 20 AA;

Query Match 33.7%; Score 34; DB 6; Length 20;

Best Local Similarity 45.5%; Pred. No. 1.3e+02;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 ISLFYTEEHHI 16

| : || |||

Db 1 IQIIYTRNHEV 11

RESULT 7

ABG62573

ID ABG62573 standard; peptide; 25 AA.

XX AC

XX AC

XX 21-AUG-2002 (first entry)

DE Eubacterial MutS1 DNA polymerase III beta subunit binding peptide #78.

XX DNA polymerase III; beta subunit; eubacteria; antibacterial;
 XX eubacterial infection.

OS Streptococcus pyogenes.

XX WO200238596-A1.

XX 16-MAY-2002.

XX 08-NOV-2001; 2001WO-AU001436.

XX 08-NOV-2000; 2000AU-00001320.

PR 06-FEB-2001; 2001AU-00002919.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX Dalrymple BP, Kongsuwan K, Wijffels GL, Jennings PA, Kemp GW;

XX WPI; 2002-471546/50.

XX New molecule having surface analogous to surface of domain of eubacterial
 PT beta protein contacted by proteins that interact with beta protein,
 PT useful to identify inhibitors of beta protein-ligand interaction.

XX Example 1; Page 32; 326pp; English.

XX The invention relates to a molecule (I) comprising a surface (S)
 CC analogous to the surface of the domain of eubacterial beta protein
 CC contacted by proteins that interact with beta protein, where the surface
 CC is defined by the residues X(170), X(172), X(175), X(177), X(241),
 CC X(242), X(247), X(346), X(360), and X(362), where the superscript numbers
 CC designate the position of residues in Escherichia coli beta protein, or
 CC the equivalent residues in homologues from other species of eubacteria,
 CC and where: X(170) = Val, Ile, Ala, Thr, Ser or Glu; X(172) = Thr, Ser or
 CC Ile; X(175) = His, Tyr, Phe, Lys, Ile, Gln or Leu; X(177) = Leu, Met,
 CC X(247) = Val, Ile, Ala, Phe, Leu or Met; X(346) = Ser, Pro, Ala, Tyr or
 CC Lys; X(360) = Ile, Leu or Val; and X(362) = Met, Leu, Val, Ser, Thr or
 CC Arg. Also included are methods of identifying a modulator of the
 CC interaction between a eubacterial beta protein and proteins that interact
 CC with them, reducing (M4) the effect of eubacterial infestation of a
 CC biological system, involves delivering to a system infested with a
 CC eubacterial species, a modulator of the interaction between eubacterial
 CC beta protein and proteins that interact with the beta protein; and (4) a

CC template (II) for the design of a compound that binds to at least part of
 CC (S) of beta protein as defined above comprises a (P) such as X¹IX²,
 CC X³X¹X², X³X¹X²X⁴, GlnX⁵X³X¹X², GlnX⁵XX⁶X³X⁶, where: x = any
 CC amino acid residue; X¹ = Leu, Met, Ile, or Phe; X² = Leu, Ile, Val,
 CC Cys, Phe, Tyr, Trp, Asp, Ala or Gly; X³ = Ala, Gly, Thr, Asn, Asp,
 CC Ser, or Pro; X⁴ = Ala or Gly; X⁵ = Leu; and X⁶ = Leu, Ile, Val, Cys,
 CC Phe, Tyr, Trp or Pro. The method are useful for identifying a modulator
 CC of the interaction between a eubacterial beta protein and proteins that
 CC interact with the beta protein. (M4) is useful for reducing the effect of
 CC eubacterial infestation of a biological system. The compounds identified
 CC using above mentioned methods are useful as antibacterial agent for
 CC treatment or prevention of disease in humans, animals and plants. The
 CC present sequence is a eubacterial peptide from a DNA binding protein or
 CC polymerase which contains a DNA polymerase III beta subunit binding site

XX SQ Sequence 25 AA;

Query Match 33.7%; Score 34; DB 5; Length 25;

Best Local Similarity 56.2%; Pred. No. 1.6e+02;

Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 3 QAKISLFYTEE--HEI 16

| : ||| || |||

Db 9 QGQLSLFGDEKAHEI 24

RESULT 8

ABR81914

ID ABR81914 standard; peptide; 28 AA.

XX AC

XX ABR81914;

XX 18-SEP-2003 (first entry)

DE Human integrin alpha 4 peptide SEQ ID NO:60.

XX Human; integrin alpha4beta1; fibronectin; integrin alpha 4; VCAM;
 KW integrin beta 1; vascular cell adhesion molecule; angiogenesis;
 KW integrin alpha4beta1 binding inhibitor; angiogenesis inhibitor;
 KW cyostatic; ophthalmological; antineumatic; antiarthritic; cancer;
 KW antiinflammatory; osteopathic; dermatological; gene therapy; psoriasis;
 KW pathological condition; diabetic retinopathy; macular degeneration;
 KW neovascularisation; rheumatoid arthritis; osteoarthritis; skin cancer;
 KW endothelial progenitor cell; wound healing.

XX Homo sapiens.

XX WO2003019136-A2.

XX 06-MAR-2003.

XX 01-AUG-2002; 2002WO-US024573.

XX 06-AUG-2001; 2001US-0310645P.

XX (REGC) UNIV CALIFORNIA.

XX Varner JA;

XX WPI; 2003-278690/27.

XX Inhibiting angiogenesis in a tissue, useful for treating cancer,
 PT arthritis, retinopathy, psoriasis, by providing a tissue and an agent
 PT that inhibits specific binding of integrin alpha-4beta-1 to an integrin
 PT alpha-4beta-1 ligand.

XX Disclosure; Page 16; 177pp; English.

XX The present invention describes a method for inhibiting angiogenesis in a
 CC tissue comprising providing a tissue and an agent that inhibits specific
 CC binding of integrin alpha4beta1 to an integrin alpha4beta1 ligand, and
 CC treating the tissue with the agent under conditions where the specific
 CC binding of the alpha4beta1 to the ligand is inhibited and a treated

CC tissue is produced. Angiogenesis in the treated tissue is inhibited. Also
 CC described: (1) inhibiting endothelial cell adhesion or migration; (2)
 CC detecting angiogenesis in a tissue; (3) screening a test compound; (4)
 CC isolating endothelial progenitor cells from a tissue; and (5) reducing
 CC symptoms associated with cancer in a subject, or a pathological condition
 CC is an ocular or skin tissue. An integrin alpha4beta1 binding inhibitor
 CC has cytosstatic, ophthalmologic, antirheumatic, antiarthritic,
 CC antiinflammatory, osteopathic and dermatological activities, and can be
 CC used in gene therapy. The methods are useful for treating cancer, and
 CC other pathological condition, such as diabetic retinopathy, macular
 CC degeneration by neovascularisation, rheumatoid arthritis, osteoarthritis,
 CC psoriasis or skin cancer. The methods are also useful in isolating
 CC endothelial progenitor cells, and in determining the mechanisms that
 CC underlie angiogenesis, development, wound healing and the function of the
 CC female reproductive system. The present sequence is used in the
 CC exemplification of the present invention

XX SQ Sequence 28 AA;
 Query Match 33.7%; Score 34; DB 6; Length 28;
 Best Local Similarity 63.6%; Pred. No. 1.8e-02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 3 QAKISLFYTEE 13
 |||||:
 Db 17 QAGISSFYTKD 27

RESULT 9
 AAB35436
 ID AAB35436 standard; peptide; 10 AA.
 AC AAB35436;
 DT 23-MAY-2001 (first entry)
 DE Nascent protein detection method related peptide #11.
 KW Nascent protein detection; protein analysis; aminoacylated tRNA;
 KW BODIPY marker; disease diagnosis.
 OS Unidentified.
 PN WO200114578-A1.
 PD 01-MAR-2001.
 PF 23-AUG-2000; 2000WO-US023233.
 PR 25-AUG-1999; 99US-00382736.
 PR 25-AUG-1999; 99US-00382950.
 PA (AMBE-) AMBERGEN INC.
 PI Rothschild KJ, Gite S, Olejnik J;
 DR WPI; 2001-168972/17.
 PT Method for detecting nascent proteins by fluorescence comprises
 PT misaminoacylating a tRNA molecule with a marker compound, useful for
 PT detecting mutations in proteins, e.g. cancer.
 PS Disclosure; Fig 2; 204pp; English.

CC The present invention describes a method of detecting nascent proteins
 CC involving aminoacylating a tRNA molecule with a 4,4-difluoro-4-bora-3A,4A
 CC -diaz-a-s-indacene (BODIPY) marker leading to the production of a
 CC misaminoacylated tRNA. This enables the detection, isolation and analysis
 CC of nascent proteins using UV without the usual accompanying radioactivity
 CC problems. It may be used to detect mutations, for example in cancer,
 CC Duchenne muscular dystrophy, adenomatous polyposis coli and colon cancer

XX SQ Sequence 10 AA;

Query Match 32.7%; Score 33; DB 4; Length 10;
 Best Local Similarity 55.6%; Pred. No. 86;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 10 YTEHEIMK 18
 ||::|||
 Db 2 YTKDHDIRK 10

RESULT 10
 ADA09802
 ID ADA09802 standard; peptide; 10 AA.
 AC ADA09802;
 DT 06-NOV-2003 (first entry)
 DE Nascent peptide #1 used to show method of translation.
 KW Non-radioactive marker; nascent protein detection method;
 KW cellular translation system; cell-free translation system;
 KW dipyrrometheneboron difluoride dye;
 KW 4,4-difluoro-4-bora-3A,4A-diaz-a-s-indacene dye; vaccine; drug;
 KW human disease screening; human disorder.
 OS Unidentified.

XX US2003092031-A1.
 XX 15-MAY-2003.
 XX 18-JUN-2002; 2002US-00174368.
 XX 25-AUG-1999; 99US-00382736.
 PR 23-AUG-2000; 2000WO-US023233.
 PR 21-JUN-2002; 2002US-00049332.
 XX (AMBE-) AMBERGEN INC.
 XX Rothschild KJ, Gite S, Olejnik J;
 DR WPI; 2003-576764/54.
 PT Detecting, analyzing or isolating nascent proteins comprises introducing
 PT a modified nucleic acid template into a cellular or cell-free translation
 PT system to generate a nascent protein having at least an N-terminal
 PT marker.

XX Disclosure; Fig 2; 76pp; English.
 CC The present invention relates to non-radioactive markers used in the
 CC detection and analysis of nascent proteins translated in cellular or cell
 CC -free translation systems. The preferred non-radioactive markers are
 CC dipyrrometheneboron difluoride (4,4-difluoro-4-bora-3A,4A-diaz-a-s-
 CC indacene) dyes. The detection method of the invention is a gel-free
 CC method that comprises introducing a modified nucleic acid template into a
 CC translation system under conditions such that a nascent protein is
 CC generated, the protein comprising at least an N-terminal marker. The
 CC method is useful in detecting, analysing and isolating nascent proteins
 CC produced in a cell-free or cellular translation system without the use of
 CC radioactive amino acids or other radioactive labels. Compositions
 CC comprising nascent proteins translated in the presence of markers may be
 CC used as vaccines or as drugs for humans and other animals. The method and
 CC a kit containing reagents for the detection of nascent proteins may be
 CC used as a rapid means to screen humans or other animals for the presence
 CC of certain diseases or disorders. The present sequence represents a
 XX nascent peptide used to show the method of translation.

XX SQ Sequence 10 AA;
 Query Match 32.7%; Score 33; DB 7; Length 10;
 Best Local Similarity 55.6%; Pred. No. 86;

Matches	5:	Conservative	3:	Mismatches	1:	Indels	0:	Gaps	0:
---------	----	--------------	----	------------	----	--------	----	------	----

```
Qy      : 10 YTEEHIMK 18
          ||::||
Db      : 2 YTKDHDIRK 10
```

RESULT 11

ADC15517	ADC15517 standard; peptide; 10 AA.
ID	

AC : ADC15517;

DT 18-DEC-2003 (first entry)

XX DE Aminoacylpolynucleotide preparation method-related peptide #1.

AA aminoacylpolynucleotide; protective group; marker; nascent protein;
KW non-radioactive labelling; detection; quantitation; isolation.
KW non-radioactive labelling; detection; quantitation; isolation.

Unidentified.

XX	Key	Location/Qualifiers
FH		

FT : Misc-difference 5

Misc-difference 6

LXXI

FM XX

XX
F7
C63-17

XX
PF
PF
17-JAN-2003; 2003MO=0300139Z.

PR 16--JAN-2003: 2003US-00349841
PR 17--JAN-2002: 2002US-00349841

XX
PA
(AMBE-) AMBERGEN INC.

XX
PI Olejnik J. Krzymanska-Olejnik E, Mamaev S, Rothschild K;

XX
DR
WPI: 2003-697532/66.

UR N-PSDB; ADU13316.
YX

Preparation of aminoacylpolynucleotide conjugate useful for introduction of markers involves attaching protective group to alpha-amino group, followed by conversion to protected amino acid ester and reaction with unprotected polynucleotide.

PS Disclosure: Fig 1: 163pp: English.

The invention comprises a method for the preparation of an aminoacylpolynucleotide that contains a protective group. The method of the invention is useful for the introduction of markers into nascent proteins for non-radioactive labelling, detection, quantitation and isolation of nascent proteins. The present amino acid sequence is used in the exemplification of the invention.

Sequence 10 AA;

Query Match	32.7%	Score 33;	DB 7;	Length 10;
Best Local Similarity	55.6%;	Pred. No. 86;		
Matches	5;	Conservative	3;	Mismatches 1;
		Indels	0;	Gaps 0;

QY 10 YTEEHEIMK 18
||::|||
Db 2 YTKDHDIRK 10

RESULT 12

ADD23457
ID: ADD23457 standard; peptide: 10 AA.

XX

ID AAY30497 standard; peptide; 15 AA.
 AC AAY30497;
 XX
 DT 15-NOV-1999 (first entry)
 XX
 DE Mutant sequence of the major epitopic region 2 of the CD2 protein.
 XX
 KW CD2 protein; rapid mutational analysis method; protein epitope mapping;
 KW binding domain mapping; binding capacity; anti-CD2 antibody;
 KW anti-CD4 antibody; ligand binding site study.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FN US595264-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 11-OCT-1994; 94US-00320663.
 XX
 PR 15-APR-1988; 88US-00181826.
 PR 27-FEB-1992; 92US-00842465.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Seed B, Peterson A;
 XX
 DR WPI; 1999-550602/46.
 XX
 PT Rapid mutational analysis method for mapping protein epitopes.
 XX
 PS Disclosure; Page; 27pp; English.
 XX
 CC AAY30496-513 represent amino acid substitution variants of a major
 CC epitopic region 2 of the CD2 protein. The peptides were acquired by
 CC mutant selection. The CD2 protein and epitopes are used to demonstrate
 CC the method of the invention. The specification describes a rapid
 CC mutational analysis method for mapping protein epitopes and binding
 CC domains, by identifying substitution mutations that result in the loss of
 CC binding capacity. The method may be used for mapping protein epitopes,
 CC antigenic domains and binding sites. It has been used for mapping binding
 CC sites for sixteen anti-CD2 and anti-CD4 monoclonal antibodies. The method
 CC is especially useful for ligand binding site studies for the design of
 CC new ligands and drugs. note: This sequence does not appear in the
 CC specification it was created using information provided
 XX
 SQ Sequence 15 AA;
 Query Match 31.7%; Score 32; DB 2; Length 15;
 Best Local Similarity 33.3%; Pred. No. 2e+02; 2; Indels 0; Gaps 0;
 Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
 QY 5 KISLFYTEHEI 16
 Db |:::|:|:
 1 KVSIVHTKEKNV 12
 RESULT 14
 AAY67695
 XX AAY67695 standard; peptide; 20 AA.
 AC
 AC AAY67695;
 XX
 DT 23-MAR-2000 (first entry)
 XX
 DE Peptide #95 for detecting hepatitis C virus infection.
 DE
 KW Hepatitis C virus; HCV; increased structural stability; E1 region;
 KW diagnostic antigen.
 XX
 OS Hepatitis C virus.
 XX

PN WO9962945-A2.
 XX
 PD 09-DEC-1999.
 XX
 PF 04-JUN-1999; 99WO-US012446.
 XX
 PR 05-JUN-1998; 98US-0088229P.
 PR 01-SEP-1998; 98US-0098705P.
 PR 15-SEP-1998; 98US-0100422P.
 PR 28-JAN-1999; 99WO-US001726.
 XX
 PA (PEPT-) PEPTIDE SOLUTIONS INC.
 XX
 PI Chowdhury MA, Bernstein D, Motsenbocker MA;
 XX
 DR WPI; 2000-086953/07.
 XX
 PT Improving properties of peptides for use as diagnostic antigens or for
 PT preventing or treating infections.
 XX
 PS Claim 55; Page 66; 83pp; English.
 XX
 CC This is a peptide from the immunoreactive region of the E1 region of
 CC hepatitis C virus (HCV). The peptide is useful for detecting HCV
 CC infection. The invention relates to peptides derived from HCV and also
 CC HIV-1 which have been modified for use as diagnostic antigens in the
 CC treatment or prevention of infection. The structural stability of the
 CC peptides can be increased in four different ways; through the replacement
 CC of a hydrophobic amino acid with a less hydrophobic amino acid; through
 CC an increase in the amount of secondary structure (i.e. alpha helix) in
 CC the peptide; through the removal of a positive charge from the peptide;
 CC or through the constraint of the epitopic sequence via the formation of a
 CC covalent crosslink. Modified peptides of the invention are used to detect
 CC infectious agents specifically HCV. Other detectable agents include HIV-1
 CC Group O viruses; human T-cell lymphotropic virus-I or -II; and the
 CC causative agent of syphilis. The peptides can be used for prevention or
 CC treatment of infections (e.g. as vaccines, or where expressed from a
 CC transgene). More generally almost any peptide can be similarly modified,
 CC e.g. cytokines or interferons; major histocompatibility complex antigens;
 CC hormones; growth factors; tumour markers or suppressors, or antigens from
 CC many other pathogens
 XX
 SQ Sequence 20 AA;
 Query Match 31.7%; Score 32; DB 3; Length 20;
 Best Local Similarity 54.5%; Pred. No. 2.7e+02;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 7 SLFYTEHEIM 17
 Db |:::|:|:
 5 SIVYETEHEIM 15
 RESULT 15
 AAR31936
 ID AAR31936 standard; protein; 25 AA.
 XX
 AC AAR31936;
 XX
 DT 25-MAR-2003 (revised)
 DT 01-JUN-1993 (first entry)
 XX
 DE Human TC-CSF peptide #2.
 XX
 KW Clone; human; T cell; colony stimulating factor; TC-CSF; AIDS; cancer;
 KW restriction site; ligand; immunoassay; hybridisation assay; primer.
 OS Homo sapiens.
 XX
 FN WO9300923-A1.
 XX
 PD 21-JAN-1993.
 XX

PF 07-JUL-1992; 92WC-US005707.
XX
PR 12-JUL-1991; 91US-00729135.
PR 19-AUG-1991; 91US-00747784.
PR 01-NOV-1991; 91US-00788115.
XX
XX (CHOU/) CHOUDHURY C.
PA
XX
XX Choudhury C;
PI
XX
XX WPI; 1993-045234/05.
XX
XX New T-lymphocyte-derived haematopoietic growth factor (TC-CSF) - for
PT diagnosing deficiency or excess of TC-CSF, stimulating haematopoiesis and
PT restoring immune function.
XX
XX Disclosure; Page 90; 107pp; English.
XX
XX The sequences given in AAR31935-36 represent partial human T cell colony
CC stimulating factor (TC-CSF) peptides. Amino acids represented by X in
CC these sequences have been identified as unknown residues. They represent
CC mutations or sequencing errors. TC-CSF polypeptides may be used in
CC pharmaceutical compositions for diagnosis and treatment of immune-
CC compromised patients, including AIDS patients and certain types of cancer
CC patients. TC-CSF and TC-CSF ligands can be used in immunoassay. The TC-
CC CSF coding sequence, or portions of it, may be used in hybridisation
CC assay or as primers. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX
SQ Sequence 25 AA;
Query Match 31.7%; Score 32; DB 2; Length 25;
Best Local Similarity 42.9%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 5 KISLFYTEEHEIMK 18
||:||||
Db 5 KINLFYNSNSPVWK 18
Search completed: July 15, 2004, 14:46:43
Job time : 42.1333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:40:36 ; Search time 8.53333 Seconds
(without alignments)
225.449 Million cell updates/sec

Title: US-09-171-432A-43
Perfect score: 101
Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 6282

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	34	33.7	30	D82251	hypothetical prote
2	31	30.7	26	A81209	hypothetical prote
3	28	27.7	25	S21204	H4-transporting tw
4	28	27.7	30	PC2251	D-tagatose 3-epime
5	27	26.7	29	S23578	H4-transporting tw
6	27	26.7	30	P00444	hypothetical prote
7	27	26.7	30	E84577	hypothetical prote
8	26	25.7	20	S17461	flavodoxin B - Azo
9	25	24.8	23	G81890	hypothetical prote
10	25	24.8	28	T06340	ribosomal protein
11	25	24.8	30	C81541	hypothetical prote
12	24	23.8	22	A28524	diaminopropionate
13	24	23.8	23	S24279	hypothetical prote
14	24	23.8	23	FN0055	unidentified QM003
15	24	23.8	28	A60359	pollen allergen DG
16	24	23.8	30	S21814	H4-exporting ATPas
17	24	23.8	30	A44913	34K core flagella
18	23	22.8	15	E41383	23K variable histo
19	23	22.8	16	S22040	cob protein - comm
20	23	22.8	21	S49040	cpb-protein 19K -
21	23	22.8	22	S52357	hypothetical prote
22	23	22.8	23	I61239	heparin-binding ep
23	23	22.8	24	A33262	heparin-binding gr
24	23	22.8	24	G64615	hypothetical prote
25	23	22.8	25	JF0064	ribosomal protein
26	23	22.8	26	A49489	pol polyprotein -
27	23	22.8	26	S49367	repA protein - Cam
28	23	22.8	30	S21816	H4-exporting ATPas
29	22	21.8	14	B83836	hypothetical prote

30	22	21.8	15	2	S42741	ubiquinol-cytochro
31	22	21.8	16	2	S5307	glutathione transf
32	22	21.8	19	2	S60110	hypothetical prote
33	22	21.8	22	2	I46682	Ca2+ ATPase, adult
34	22	21.8	22	2	F84018	hypothetical prote
35	22	21.8	25	2	PC1119	atrolysin A (EC 3.
36	22	21.8	26	2	S68632	iota toxin - Clost
37	22	21.8	26	2	S49370	repA protein - Cam
38	22	21.8	26	2	A40630	hypothetical prote
39	22	21.8	26	2	S60884	RfBC protein - Kle
40	22	21.8	27	2	PH1371	Ig heavy chain DJ
41	22	21.8	27	2	B47295	homeodomain protei
42	22	21.8	27	2	T06931	hypothetical prote
43	22	21.8	28	2	I58115	cystic fibrosis tr
44	22	21.8	30	2	A34461	heat shock protein
45	22	21.8	30	2	F89406	protein R10E8.7 [i

ALIGNMENTS

RESULT 1

D82251
hypothetical protein VC1030 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82251
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82251
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-30 <HEI>
A:Cross-references: GB:AE004184; GB:AE003852; NID:g9655483; PIDN:AAF94189.1; GSFDDB:GN001:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1030
A:Map position: 1

Query Match 33.7%; Score 34; DB 2; Length 30;
Best Local Similarity 43.8%; Pred. No. 40;
Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 5 KISLFYTEHEIMKFS 20
| | | | | | | | | | | | | | | | | | | | | |
Db 9 KFILFY--EHELLRLN 22

RESULT 2

A81209
hypothetical protein NMB0362 [imported] - Neisseria meningitidis (strain MC58 serogroup F
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: A81209
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: A81209
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-26 <TET>

A:Cross-references: GB:AE002392; GB:AE002098; NID:g7225573; PIDN:AAF40805.1; PID:g722558:
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0362

Query Match 30.7%; Score 31; DB 2; Length 26;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAKISLFYT 11
 :||:||||
 Db 4 KAVISLFYT 12

RESULT 3
 S21204
 H+-transporting two-sector ATPase (EC 3.6.3.14) alpha chain - spinach mitochondrion (fragment)
 C;Species: mitochondrion Spinacia oleracea (spinach)
 C;Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 03-Jun-2002
 C;Accession: S21204
 R;Namasur, B.; Glaser, E.
 Euz. J. Biochem. 205, 409-416, 1992
 A;Title: Plant mitochondrial F(0)F(1) ATP synthase. Identification of the individual subunits
 A;Reference number: S21204; MUID:92209531; PMID:1313368
 A;Accession: S21204
 A;Molecule type: protein
 A;Residues: 1-25 <HAW>
 C;Genetics:
 A;Genome: mitochondrion
 C;Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase all subunits
 C;Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion

Query Match 27.7%; Score 28; DB 2; Length 25;
 Best Local Similarity 55.6%; Pred. No. 3.3e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAKISLFYT 11
 :||:||||
 Db 15 ESRISNFT 23

RESULT 4
 PC2251
 D-tagatose 3-epimerase (EC 5.1.3.-) - Pseudomonas sp. (fragment)
 C;Species: Pseudomonas sp.
 C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 24-Feb-1995
 C;Accession: PC2251
 R;Itoh, H.; Okaya, H.; Khan, A.R.; Tajima, S.; Hayakawa, S.; Izumori, K.
 Biosci. Biotechnol. Biochem. 58, 2168-2171, 1994
 A;Title: Purification and characterization of D-tagatose 3-epimerase from Pseudomonas sp.
 A;Reference number: PC2251
 A;Accession: PC2251
 A;Molecule type: protein
 A;Residues: 1-30 <ITO>
 A;Experimental source: ST-24
 C;Comment: This enzyme catalyzes the epimerization of free keto-sugars.
 C;Keywords: isomerase

Query Match 27.7%; Score 28; DB 2; Length 30;
 Best Local Similarity 57.1%; Pred. No. 4.1e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 KISLFYT 11
 |:||||
 Db 3 KVGMYFT 9

RESULT 5
 S04578
 H+-transporting two-sector ATPase (EC 3.6.3.14) beta chain - spinach chloroplast (fragment)
 C;Species: chloroplast Spinacia oleracea (spinach)
 C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002
 C;Accession: S02578
 R;Jue, Z.; Miller, C.G.; Zhou, J.M.; Boyer, P.D.
 FEBS Lett. 223, 391-394, 1987
 A;Title: Catalytic and noncatalytic nucleotide binding sites of chloroplast F(1) ATPase.
 A;Reference number: S02578; MUID:88030083; PMID:2889622

A;Accession: S02578
 A;Molecule type: protein
 A;Residues: 1-18,19-29 <XUB>
 C;Genetics:
 A;Genome: chloroplast
 C;Superfamily: H+-transporting ATP synthase alpha chain; H+-transporting ATP synthase all subunits
 C;Keywords: ATP biosynthesis; chloroplast; hydrolase; thylakoid

Query Match 26.7%; Score 27; DB 2; Length 29;
 Best Local Similarity 100.0%; Pred. No. 5.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 EEHEI 16
 :||:||||
 Db 22 EEHEI 26

RESULT 6
 PQ0444
 hypothetical protein 30 - slime mold (Dictyostelium giganteum) plasmid Dgpl (fragment)
 C;Species: Dictyostelium giganteum
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 29-Oct-1999
 C;Accession: PQ0444
 R;Yin, Y.; Welker, D.L.
 Plasmid 28, 37-45, 1992
 A;Title: Dictyostelium giganteum plasmid Dgpl is a member of the Ddp2 plasmid family.
 A;Reference number: PQ0444; MUID:92390516; PMID:1518911
 A;Accession: PQ0444
 A;Molecule type: DNA
 A;Residues: 1-30 <YIN>
 A;Cross-references: GB:S43953; NID:g254926; PIDN:AAB23143.1; PID:g254927
 A;Experimental source: strain DG61
 C;Genetics:
 A;Genome: plasmid

Query Match 26.7%; Score 27; DB 2; Length 30;
 Best Local Similarity 40.0%; Pred. No. 6.1e+02;
 Matches 6; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 KISLFYTEHEIMKF 19
 |:||||
 Db 9 KVGFLYHCRSEISKF 23

RESULT 7
 E64577
 hypothetical protein HP0461 - Helicobacter pylori (strain 26695)
 C;Species: Helicobacter pylori
 C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
 C;Accession: E64577
 R;Tomb, J.P.; White, O.; Karlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
 Nature 388, 539-547, 1997
 A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A;Reference number: A64520; MUID:97394467; PMID:9252185
 A;Accession: E64577
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-30 <TOM>
 A;Cross-references: GB:AE000561; GB:AE000511; NID:g2313564; PIDN:AAD07538.1; PID:g2313564

Query Match 26.7%; Score 27; DB 2; Length 30;
 Best Local Similarity 33.3%; Pred. No. 6.1e+02;
 Matches 5; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEE 15
 :||:||||
 Db 15 FKQQLLTFFYKTRHK 29

RESULT 8

C;Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 24.8%; Score 25; DB 2; Length 28;
Best Local Similarity 40.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFVTEEHE 15
|::|||::|||
Db 3 FIESLSQMFDEERE 17
|::|||::|||

RESULT 11

C81541
hypothetical protein CP0767 [imported] - Chlamydoiphila pneumoniae (strain AR39)
C:Species: Chlamydoiphila pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C:Accession: C81541
C:; Dodson, R.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
R.; Read, T.D.; Gwinm, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: C81541
A:Status: preliminary
A:; Molecule type: DNA
A:Residues: 1-30 <REA>
A:Cross-references: GB:AE002236; GB:AE002161; NID:g7189675; PIDN:AAF38567.1; PID:g7189686
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0767

Query Match 24.8%; Score 25; DB 2; Length 30;
Best Local Similarity 40.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 ISLFVTEEHE 15
:|::|||::|
Db 16 VALTYVFEHD 25
:|::|||::|

RESULT 12

A28524
diaminopropionate ammonia-lyase - Salmonella typhimurium (fragment)
N:Alternate names: diaminopropionatase
C:Species: Salmonella typhimurium
C:Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 18-Jun-1993
C:Accession: A28524
R:Nagasawa, T.; Tanizawa, K.; Satoda, T.; Yamada, H.
J. Biol. Chem. 263, 958-964, 1988
A:Title: Diaminopropionate ammonia-lyase from Salmonella typhimurium. Purification and
tide.
A:Reference number: A28524; MUID:88087224; PMID:3275662
A:Accession: A28524
A:; Molecule type: protein
A:Residues: 1-22 <NAG>

Query Match 23.8%; Score 24; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 14 HEIMKF 19
|::|||::|
Db 2 HELIKY 7
|::|||::|

RESULT 13

S24279
hypothetical protein 3 - porcine respiratory virus (strain 86/137004) (fragment)
C:Species: porcine respiratory virus
A:Variety: strain 86/137004
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
C:Accession: S24279; S21308
R:Britton, P.; Mawditt, K.L.; Page, K.W.

Virus Res. 21, 181-198, 1991
A; Title: The cloning and sequencing of the virion protein genes from a British isolate c
A; Reference number: S24279; MUID:92116634; PMID:1662846
A; Accession: S24279
A; Molecule type: genomic RNA
A; Residues: 1-23

A; Cross-references: EMBL:X60056; NID:961342; PIDN:CAA42654.1; PID:g61343
A; Experimental source: strain 86/137004
R; Britton, P.; Mawditt, K.L.
submitted to the EMBL Data Library, September 1990
A; Description: Nucleotide sequence of the membrane protein and ORF-4 genes from porcine
A; Reference number: S21308
A; Accession: S21308
A; Molecule type: mRNA
A; Residues: 1-23

A; Cross-references: EMBL:X55980; NID:961337; PIDN:CAA39450.1; PID:g61338

Query Match 23.8%; Score 24; DB 2; Length 23;
Best Local Similarity 40.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTERHE 15
|||||:|
Db 1 FSQAVLNEIDLKEE 15

RESULT 14
PN0055
unidentified QM0037 protein - mouse (fragments)
C; Species: Mus musculus (house mouse)
C; Date: 29-Oct-1997 #sequence_revision 29-Oct-1997 #text_change 23-Jan-1998
C; Accession: PN0055
R; Kato, H.
Kawasaki Igakkaishi 22, 245-259, 1996
A; Title: Analysis of proteins isolated by two dimensional electrophoresis of mouse neuro
A; Reference number: PN0041
A; Accession: PN0055
A; Molecule type: protein
A; Residues: 1-23 <KAT>
A; Experimental source: neuroblastoma cell
C; Comment: The molecular mass is 13,500 and the pI is 5.64.
C; Keywords: brain

Query Match 23.8%; Score 24; DB 2; Length 23;
Best Local Similarity 56.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 LFYTEE 13
:|
Db 11 MFYPEE 16

RESULT 15
A60359
pollen allergen DG3 - orchard grass (fragment)
C; Species: Dactylis glomerata (orchard grass)
C; Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 03-Nov-2000
C; Accession: A60359
R; Walsh, D.J.; Matthews, J.A.; Denmeade, R.; Maxwell, P.; Davidson, M.; Walker, M.R.
Int. Arch. Allergy Appl. Immunol. 91, 419-425, 1990
A; Title: Monoclonal antibodies to proteins from cocksfoot grass (Dactylis glomerata) pol
A; Reference number: A60359; MUID:91007970; PMID:2210876
A; Accession: A60359
A; Molecule type: protein
A; Residues: 1-28 <WAL>
C; Superfamily: grass pollen allergen IX
C; Keywords: pollen

Query Match 23.8%; Score 24; DB 2; Length 28;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 TEEHEIMK 18

||| :|
Db 16 TEEQKLME 23

Search completed: July 15, 2004, 14:52:11
Job time : 9.53333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:01 ; Search time 5.3333 Seconds

(without alignments)
195.263 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101

Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 2073

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	27.7	20	1	CATA_ACIRA
2	28	27.7	25	1	ATP0_SPIOL
3	25	24.8	18	1	RL24_SERMA
4	24	23.8	30	1	VAA1_EQUAR
5	23	22.8	7	1	FAR2_ASCSU
6	23	22.8	19	1	P5BN_SYNNU
7	23	22.8	24	1	HPTA_RABIT
8	23	22.8	30	1	VAA1_PSIU
9	23	22.8	30	1	VAA2_PSIU
10	22	21.8	16	1	BBB_BASAL
11	22	21.8	20	1	FLAW_AZOVI
12	22	21.8	23	1	P53_ESEPD
13	22	21.8	23	1	PS4_PSEPD
14	22	21.8	27	1	YCX7_CYAPA
15	21	20.8	9	1	FAR1_CALVO
16	21	20.8	18	1	RL24_PROVU
17	21	20.8	19	1	SCX6_TITBA
18	21	20.8	19	1	TRPB_KLEAE
19	21	20.8	20	1	EFTU_MYCSY
20	21	20.8	21	1	FIBB_CEREL
21	21	20.8	21	1	MISG_MISAN
22	21	20.8	24	1	PEDG_AMEYE
23	21	20.8	29	1	28KD_TRIPO
24	20	19.8	10	1	TKN1_SCYCA
25	20	19.8	14	1	FIBA_HORSE
26	20	19.8	16	1	UPAB_HUMAN
27	20	19.8	20	1	ACPH_BOVIN
28	20	19.8	22	1	CLH1_CHEAL
29	20	19.8	22	1	HCY4_HORAM
30	19	18.8	9	1	FAR2_CALVO
31	19	18.8	9	1	FARA_CALVO
32	19	18.8	13	1	MLA_CAMDR
33	19	18.8	15	1	OBPA_MAMBR

ALIGNMENTS

RESULT 1

ID	CATA_ACIRA	STANDARD;	PRT;	20 AA.
AC	P81422;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Catechol 1,2-dioxygenase (EC 1.13.11.1) (1,2-CTD) (Fragment).			
OS	Acinetobacter radioresistens.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;			
OC	Moraxellaceae; Acinetobacter.			
OX	NCBI_TaxID=40216;			
RN	[1]			
RP	SEQUENCE			
RX	MEDLINE=98034166; PubMed=9369233;			
RA	Briganti F., Pessione E., Giunta C., Scozzafava A.;			
RT	"Purification, biochemical properties and substrate specificity of a catechol 1,2-dioxygenase from a phenol degrading Acinetobacter radioresistens.";			
RL	FEBS Lett. 416:61-64(1997).			
CC	-!- CATALYTIC ACTIVITY: Catechol + O(2) = cis,cis-muconate.			
CC	-!- COFACTOR: Binds 1 iron (ferric) ion per subunit.			
CC	-!- PATHWAY: Degradation of catechol to succinate and acetyl-CoA in the beta-ketoadipate pathway; first step.			
CC	-!- SUBUNIT: Homodimer which dissociates into active monomeric subunits at high ionic strengths			
CC	-!- SIMILARITY: Belongs to the intradiol ring-cleavage dioxygenase family.			
DR	InterPro; IPR00627; Dioxygenase.			
DR	PROSITE; PS00083; INTRADIOL DIOXYGENAS; PARTIAL.			
KW	Aromatic hydrocarbons catabolism; Oxidoreductase; Dioxygenase; Iron.			
FT	NON TER 20 20			
SQ	SEQUENCE 20 AA; 2279 MW; 7089A5038F802327 CRC64;			

Query Match 27.7%; Score 28; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 AKISLFYTEE 13
| : : | | |
Db 3 ANVKIFNTEE 12

RESULT 2

ID	ATP0_SPIOL	STANDARD;	PRT;	25 AA.
AC	P80082;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	ATP synthase alpha chain, mitochondrial (EC 3.6.3.14) (Fragment).			
DE	ATPA.			
OS	Spinacia oleracea (Spinach).			
OG	Mitochondrion.			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;			

```

OC OC Spermatozoa; Magnoliophyta; eudicotyledons; core eudicots;
OC OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RX STRAIN=cv. Medania; TISSUE=Leaf mesophyll;
RX MEDLINE=92209531; PubMed=1313368;
RA Hamasur B., Glaser E.;
RT "Plant mitochondrial F0F1 ATP synthase. Identification of the
RT individual subunits and properties of the purified spinach leaf
RT mitochondrial ATP synthase."
RL Eur. J. Biochem. 205:409-416(1992).
CC -!- FUNCTION: Produces ATP from ADP in the presence of a proton
CC gradient across the membrane. The alpha chain is a regulatory
CC subunit.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: a, b and c.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
DR PIR; S21204; S21204.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR PROSITE; PS00152; ATPASE ALPHA BETA; PARTIAL.
KW ATP synthetase; CF(1); Hydrogen ion transport;
KW Hydrolyase; ATP-binding; Mitochondrion.
FT NON TER 25 25
SQ SEQUENCE 25 AA; 2904 MW; 1B1486EBBD1A650D CRC64;

Query Match 27.7%; Score 28; DB 1; Length 25;
Best Local Similarity 55.8%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAKISLFYT 11
DB ::::|
15 ESRISNYFT 23

RESULT 3
RL24 SERMA STANDARD; PRT; 18 AA.
AC P49624;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 50S ribosomal protein L24 (Fragment).
GN RPLX.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89125589; PubMed=2464692;
RA Cerretti D.P., Mattheakis L.C., Kearney K.R., Vu L., Nomura M.;
RT "Translational regulation of the spc operon in Escherichia coli.
RT Identification and structural analysis of the target site for S8
RT repressor protein."
RL J. Mol. Biol. 204:309-329(1998).
CC -!- FUNCTION: This protein is found in the ribonucleoprotein core and
CC is involved in the early assembly of the 50S subunit. It is not
CC involved in the functions of the mature 50S subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the L24P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56983; CAA40301.1; -.
CC PIR; S21814; S21814.
CC InterPro; IPR000194; ATPase_a/bcentre.
CC PROSITE; PS00152; ATPASE ALPHA BETA; PARTIAL.
KW ATP synthetase; Hydrogen ion transport; Hydrolyase; ATP-binding;
KW Multigene family.
FT NON TER 1 1
FT NON TER 30 30
SQ SEQUENCE 30 AA; 3337 MW; 9627ED62068D761F CRC64;

Query Match 23.8%; Score 24; DB 1; Length 30;
Best Local Similarity 30.0%; Pred. No. 1.3e+03;

```

```

CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M36265; AAA26580.1; -.
CC InterPro; IPR005825; Ribosomal_L24_26.
CC PROSITE; PS01108; RIBOSOMAL_L24; PARTIAL.
KW Ribosomal protein.
FT NON TER 1 1
SQ SEQUENCE 18 AA; 2174 MW; 43606F53D34BA2E2 CRC64;

Query Match 24.8%; Score 25; DB 1; Length 18;
Best Local Similarity 33.3%; Pred. No. 5e+02;
Matches 6; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMK 18
DB |||||:|
1 FEDAKKVRFFKNSSETIK 18

RESULT 4
VAAL EQUAR STANDARD; PRT; 30 AA.
AC Q04236;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Vacuolar ATP synthase catalytic subunit A, isoform 1 (EC 3.6.3.14)
DE (Fragment).
OS Equisetum arvense (Field horsetail) (Common horsetail).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Moniliformophyceae; Equisetophyta; Sphenopsida; Equisetales;
OC Equisetaceae; Equisetum.
OX NCBI_TaxID=3258;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93138084; PubMed=8422915;
RA Starke T., Gogarten J.P.;
RT "A conserved intron in the V-ATPase A subunit genes of plants and
RT algae."
RL FEBS Lett. 315:252-258(1993).
CC -!- FUNCTION: Catalytic subunit of the peripheral V1 complex of a
CC vacuolar ATPase. V-ATPase vacuolar ATPase is responsible for
CC acidifying a variety of intracellular compartments in eukaryotic
CC cells.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a
CC peripheral catalytic V1 complex (main components: subunits A, B,
CC C, D, E, and F) attached to an integral membrane V0 proton pore
CC complex (main component: the proteolipid protein).
CC -!- MISCELLANEOUS: TWO SEPARATE GENES ENCODE THE CATALYTIC 70 kDa
CC V-ATPASE SUBUNIT IN PSILOTUM AND EUISETUM.
CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56983; CAA40301.1; -.
CC PIR; S21814; S21814.
CC InterPro; IPR000194; ATPase_a/bcentre.
CC PROSITE; PS00152; ATPASE ALPHA BETA; PARTIAL.
KW ATP synthetase; Hydrogen ion transport; Hydrolyase; ATP-binding;
KW Multigene family.
FT NON TER 1 1
FT NON TER 30 30
SQ SEQUENCE 30 AA; 3337 MW; 9627ED62068D761F CRC64;

Query Match 23.8%; Score 24; DB 1; Length 30;
Best Local Similarity 30.0%; Pred. No. 1.3e+03;

```


PP SEQUENCE.
RX MEDLINE=89338735; PubMed=2503398;
RA Ikeuchi M., Koike H., Inoue Y.;
RT N-terminal sequencing of low-molecular-mass components in
RT cyanobacterial photosystem II core complex. Two components correspond
RT to unidentified open reading frames of plant chloroplast DNA.;
RL FEBS Lett. 253:178-182(1989).

RESULT 8	VAAL_PSIU	VAAL_PSIU	STANDARD	PRT	30 AA.
ID	VAAL_PSIU	VAAL_PSIU	STANDARD	PRT	30 AA.
AC	Q04237				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE	Vacuolar ATP synthase catalytic subunit A, isoform 1 (EC 3.6.3.14) (fragment).				
DE	Psilotum nudum (Whisk fern).				
OS	Psilotum nudum (Whisk fern).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Monilliformes; Psilotophyta; Psilotales; Psilotaceae; Psilotum.				
OX	NCBI_TaxID=3240;				
OX	[1]_TaxID=3240;				
RN	SEQUENCE FROM N.A.				
RP					
RA	MEDLINE=93138084; PubMed=8422915;				
RX	Stärke T., Gogarten J.P.;				

DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Flavodoxin 1 (Fragment).
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Azotobacter.
 OX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=OP / UW136;
 RX MEDLINE=96276406; PubMed=8694750;
 RA Gangawaran R., Eady R.R.;
 RT "Flavodoxin 1 of Azotobacter vinelandii: characterization and role in
 electron donation to purified assimilatory nitrate reductase."
 RL Biochem. J. 317:103-108(1996).
 CC -!- FUNCTION: Low-potential electron donor to a number of redox
 enzymes.
 CC -!- COFACTOR: FMN.
 CC -!- SUBUNIT: Monomer.
 CC -!- MASS SPECTROMETRY: MW=19430; MW_ERR=3; METHOD=Electrospray.
 CC -!- SIMILARITY: Belongs to the flavodoxin family.
 CC -!- SIMILARITY: Contains 1 flavodoxin-like domain.
 DR InterPro; IPR008254; Flav nitox synth.
 DR InterPro; IPR001226; Flavodoxin.
 DR PROSITE; PS00201; FLAVODOXIN; PARTIAL.
 DR PROSITE; PS00902; FLAVODOXIN LIKE; 1.
 KW Electron transport; Flavoprotein; FMN.
 FT DOMAIN 3 ->20 FLAVODOXIN-LIKE.
 FT NON TER 20 20
 SQ SEQUENCE 20 AA; 2043 MW; 95B3601FAD7D4A36 CRC64;

Query Match 21.8%; Score 22; DB 1; Length 20;
 Best Local Similarity 42.9%; Pred. No. 1.7e+03;
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 AKISLFY 10
 :.:|:|
 DB 1 SRIGIFY 7

RESULT 12
 PS3_PSEPD
 ID PS3_PSEPD STANDARD; PRT; 23 AA.
 AC P83190;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pseudin 3.
 OS Pseudis paradoxa (Paradoxical frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaidea; Pseudidae;
 OC Pseudis.
 OX NCBI_TaxID=43558;
 RN [1]
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
 RC TISSUE=Skin secretion;
 RX MEDLINE=21547799; PubMed=11689009;
 RA Olson L. III, Soto A.M., Knoop F.C., Conlon J.M.;
 RT "Pseudin-2: an antimicrobial peptide with low hemolytic activity from
 the skin of the paradoxical frog."
 RL Biochem. Biophys. Res. Commun. 288:1001-1005(2001).
 CC -!- FUNCTION: Possesses antifungal activity against C.albicans and is
 also active against E.coli and S.aureus.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 CC -!- MASS SPECTROMETRY: MW=2569.6; METHOD=Electrospray.
 DR GO; GO:0003799; F:antifungal peptide activity; IDA.
 KW Amphibian defense peptide; Fungicide; Antibiotic.
 SQ SEQUENCE 23 AA; 2570 MW; D9FAAB75B37D2373 CRC64;

Query Match 21.8%; Score 22; DB 1; Length 23;
 Best Local Similarity 60.0%; Pred. No. 2e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 14 HEIMK 18
 :.:|:|
 DB 13 HEVIK 17

RESULT 13
 PS4_PSEPD
 ID PS4_PSEPD STANDARD; PRT; 23 AA.
 AC P83191;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pseudin 4.
 OS Pseudis paradoxa (Paradoxical frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaidea; Pseudidae;
 OC Pseudis.
 OX NCBI_TaxID=43558;
 RN [1]
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
 RC TISSUE=Skin;
 RX MEDLINE=21547799; PubMed=11689009;
 RA Olson L. III, Soto A.M., Knoop F.C., Conlon J.M.;
 RT "Pseudin-2: an antimicrobial peptide with low hemolytic activity from
 the skin of the paradoxical frog."
 RL Biochem. Biophys. Res. Commun. 288:1001-1005(2001).
 CC -!- FUNCTION: Possesses antifungal activity against C.albicans and is
 also active against E.coli and S.aureus.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Skin.
 CC -!- MASS SPECTROMETRY: MW=2512; METHOD=Electrospray.
 DR GO; GO:0003799; F:antifungal peptide activity; IDA.
 KW Amphibian defense peptide; Fungicide; Antibiotic.
 SQ SEQUENCE 23 AA; 2512 MW; DF3AAB75B37D2373 CRC64;

Query Match 21.8%; Score 22; DB 1; Length 23;
 Best Local Similarity 60.0%; Pred. No. 2e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 14 HEIMK 18
 :.:|:|
 DB 13 HEVIK 17

RESULT 14
 YCX7_CYAPA
 ID YCX7_CYAPA STANDARD; PRT; 27 AA.
 AC P48328;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical 3.4 kDa protein in ATPE-PETA intergenic region (OREF27).
 OS Cyanophora paradoxa.
 OG Cyanelle.
 OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
 OX NCBI_TaxID=2762;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UTEX LB 555 / Pringsheim;
 RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
 RA Bryant D.A.;
 RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa."
 RL Plant Mol. Biol. Rep. 13:327-332(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UTEX LB 555 / Pringsheim;
 RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
 RA Farley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,
 RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
 RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
 the genetic complexity of a primitive plastid."
 RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,

```

RL Schwenmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL (1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; U30821; AAA81274.1; -.
DR PIR; T06931; T06931.
DR KW Hypothetical protein; Cyanelle.
SQ SEQUENCE 27 AA; 3442 MW; F195AE2F414C3288 CRC64;

Query Match      21.8%; Score 22; DB 1; Length 27;
Best Local Similarity 16.7%; Pred. No. 2.4e+03;
Matches 2; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY      8 LFYTEEHEIMKF 19
      :|:|:|:|:|:|
DB      3 IYYLQSEKVLFF 14

RESULT 15
FAM1 CALVO
ID 1 FARM1 CALVO STANDARD; PRT; 9 AA.
AC P41856;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CalliFMRPamide 1.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliFMRPamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -!- FUNCTION: Able to induce fluid secretion from the isolated
CC salivary gland of Calliphora.
CC -!- SIMILARITY: Belongs to the FARP (FMRPamide related peptide)
CC family.
DR PIR; A41978; A41978.
DR KW Neuropeptide; Amidation.
FT MOD RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 1169 MW; 29D00699CAB6C6C7 CRC64;

Query Match      20.8%; Score 21; DB 1; Length 9;
Best Local Similarity 33.3%; Pred. No. 1.4e+05;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      11 TEEHEIMKF 19
      |:|:|:|:|
DB      1 TPQQDFMRF 9

```

Search completed: July 15, 2004, 14:47:32
Job time : 6.33333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:31 ; Search time 25.7333 Seconds
(without alignments)
245.221 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101
Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 18020

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	33.7	30	16	Q9TKG75
2	32	31.7	28	4	Q96JA8
3	31	30.7	17	1	Q9V2X3
4	31	30.7	26	16	Q9K131
5	30	29.7	17	4	Q9UCF0
6	29	28.7	29	6	Q86SH8
7	29	28.7	30	8	Q9TRV9
8	29	28.7	30	15	Q36617
9	28	27.7	30	15	Q36619
10	27	26.7	15	5	P82207
11	27	26.7	22	3	O14394
12	27	26.7	30	16	O25208
13	26.5	26.2	21	6	Q9TRV3
14	26	25.7	13	5	Q812E2
15	26	25.7	20	2	Q9R5U5
16	26	25.7	23	4	Q96HR1

17	26	25.7	23	8	Q9TKG7	Q9tkg7 lambia anta
18	26	25.7	23	8	Q9T4C9	Q9t4c9 bryopsis sp
19	26	25.7	23	8	Q9T3F1	Q9t3f1 bryopsis sp
20	26	25.7	23	8	Q9T3F3	Q9t3f3 bryopsis sp
21	26	25.7	23	8	Q9T3F2	Q9t3f2 bryopsis sp
22	26	25.7	23	8	Q9T4Q8	Q9t4q8 bryopsis sp
23	26	25.7	23	8	Q9T3S4	Q9t3s4 bryopsis sp
24	26	25.7	24	11	Q925A5	Q925a5 rattus norv
25	26	25.7	27	2	Q84DW7	Q84dw7 tropheryma
26	26	25.7	27	12	Q8QK48	Q8qk48 hepatitis c
27	26	25.7	29	16	Q87FL3	Q87fl3 vibrio para
28	26	25.7	30	15	Q36618	Q36618 ovine lenti
29	25	24.8	15	13	Q9PR29	Q9prz9 micropogoni
30	25	24.8	20	11	Q924T0	Q924t0 rattus norv
31	25	24.8	22	4	Q8NSB1	Q8nsb1 homo sapien
32	25	24.8	22	10	O22097	O22097 pharbitis n
33	25	24.8	23	16	Q9JUN1	Q9jun1 neisseria m
34	25	24.8	26	1	Q9UWG7	Q9uwg7 sulfolobus
35	25	24.8	26	2	Q9R4R1	Q9r4r1 citrobacter
36	25	24.8	27	10	Q93YC7	Q93yc7 nicotiana t
37	25	24.8	28	6	Q9TRV2	Q9trv2 canis famil
38	25	24.8	28	8	Q32307	Q32307 glycine max
39	25	24.8	29	11	O88213	O88213 mus musculu
40	25	24.8	30	15	Q36620	Q36620 ovine lenti
41	25	24.8	30	16	Q9K122	Q9k122 chlamydia p
42	24.5	24.3	26	10	O41052	O41052 porphyra sp
43	24	23.8	12	11	Q7TMM4	Q7tnm4 rattus norv
44	24	23.8	14	6	Q8HYM2	Q8hym2 felis silve
45	24	23.8	20	2	Q9R4J6	Q9r4j6 pseudomonas

ALIGNMENTS

RESULT 1

ID Q9KT75 PRELIMINARY; PRT; 30 AA.

AC Q9KT75;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein VC1030.
GN VC1030.

OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RI Nature 406:477-483(2000).
DR EMBL; AE004184; AAF94189.1; -.
DR PIR; D82251; D82251.
DR TIGR; VC1030; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 30 AA; 3699 MW; AA58B738A08449C0 CRC64;

Query Match 33.7%; Score 34; DB 16; Length 30;
Best Local Similarity 43.8%; Pred. No. 2e+02; 3; Indels 2; Gaps 1;
Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 5 KISLFYTERHEIMKFS 20
| ||| ||::: |

Db 9 KFILFY--EBELLRLN 22

RESULT 2

Q96JAB PRELIMINARY; PRT; 28 AA.
 AC Q96JAB; MEDLINE=21293069; PubMed=11283018;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Bcl-XL-binding protein c21 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Hammond P.W., Alpin J., Rise C.E., Wright M., Kreider B.L.;
 RT "In Vitro Selection and Characterization of Bcl-XL-binding Proteins
 from a Mix of Tissue-specific mRNA Display Libraries.";
 RL J. Biol. Chem. 276:20898-20906 (2001).
 DR EMBL; AF357525; AAK60629.1; -.
 FT NON TER 1 28
 FT 1
 FT 28
 SQ SEQUENCE 28 AA; 3201 MW; A7DCE42AA280631E CRC64;

Query Match 31.7%; Score 32; DB 4; Length 28;

Best Local Similarity 45.5%; Pred. No. 4e+02;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 KISLFYTEE 15

Db 16 KLKIFYTSKE 26

RESULT 3

Q9V2X3 PRELIMINARY; PRT; 17 AA.
 AC Q9V2X3; MEDLINE=88198019; PubMed=2834336;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE RNA polymerase subunit B', (Fragment).
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=197420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schallenberg J., Moes M., Truss M., Reiser W., Thomm M., Stetter K.O.,
 RA Klein A.;
 RT "Cloning and physical mapping of RNA polymerase genes from
 Methanobacterium thermoautotrophicum and comparison of homologies and
 gene orders with those of RNA polymerase genes from other Methanogenic
 archaeobacteria.";
 RL J. Bacteriol. 170:2247-2253 (1988).
 DR EMBL; M20391; AAA72654.1; -.
 DR InterPro; IPR007646; RNA_pol_Rpb2_4.
 DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
 FT NON TER 17 17
 FT 17
 FT 17
 SQ SEQUENCE 17 AA; 2137 MW; 0C03CD9E6D780560 CRC64;

Query Match 30.7%; Score 31; DB 1; Length 17;

Best Local Similarity 54.5%; Pred. No. 3.5e+02;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 10 YTEEHEIMKFS 20

Db 6 YPENHEIYIT 16

RESULT 4

Q9K131 PRELIMINARY; PRT; 26 AA.
 AC Q9K131; MEDLINE=20175755; PubMed=10710307;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein NMB0362.
 GN NMB0362.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Massignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58.";
 RL Science 287:1809-1815 (2000).
 DR EMBL; AE002392; AAF40805.1; -.
 DR PIR; A81209; A81209.
 DR TIGR; NMB0362; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 26 AA; 3101 MW; 3A4A3E475B836979 CRC64;

Query Match 30.7%; Score 31; DB 16; Length 26;

Best Local Similarity 66.7%; Pred. No. 5.4e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QAKISLFYT 11

Db 4 KAVISIFYT 12

RESULT 5

Q9UCF0 PRELIMINARY; PRT; 17 AA.
 AC Q9UCF0; MEDLINE=93309464; PubMed=8321232;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE D(TTAGGG)N-binding protein B50=NUCLEOLIN homolog.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RA Ishikawa F., Matunis M.J., Dreyfuss G., Cech T.R.;
 RT "Nuclear proteins that bind the pre-mRNA 3' splice site sequence
 r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n.";
 RL Mol. Cell. Biol. 13:4301-4310 (1993).
 SQ SEQUENCE 17 AA; 2022 MW; 03AB11FCE775E2F6 CRC64;

Query Match 29.7%; Score 30; DB 4; Length 17;

Best Local Similarity 75.0%; Pred. No. 5.2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 ISLFYTEE 13

Db 2 ISLYTGE 9

```

RESULT 6
Q865H8 PRELIMINARY; PRT; 29 AA.
AC Q865H8:
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transient receptor potential channel 2 (Fragment).
GN TRPC2.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22531908; PubMed=12631698;
RA Liman E.R., Innan H.;
RT "Relaxed selective pressure on an essential component of pheromone
transduction in primate evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:3328-3332 (2003).
DR EMBL: AY231391; AA047822.1; -.
DR GO: GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON TER 1
FT NON TER 29
SQ SEQUENCE 29 AA; 3266 MW; 9921F33BEFE7D2F7 CRC64;

Query Match 28.7%; Score 29; DB 6; Length 29;
Best Local Similarity 46.2%; Pred. No. 1.3e+03;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEE 13
|:|:|:|:|:|
Db 1 FASKICLSYFRE 13

RESULT 7
Q9T2V9 PRELIMINARY; PRT; 30 AA.
AC Q9T2V9:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VICTORIN binding protein (Fragment).
OS Cochliobolus victorinae (Drechslera victorinae).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Pleosporaceae; Cochliobolus.
OX NCBI_TaxID=40125;
RN [1]
RP SEQUENCE.
RX MEDLINE=95290998; PubMed=7773018;
RA Navarre D.A., Wolpert T.J.;
RT "Inhibition of the glycine decarboxylase multienzyme complex by the
host-selective toxin victorin.";
RL Plant Cell 7:463-471 (1995).
DR HSSP; P16048; 1HTP.
DR GO: GO:0005960; C:glycine cleavage complex; IEA.
DR GO: GO:0006546; P:glycine catabolism; IEA.
DR InterPro; IPR002930; GCV_H.
DR Pfam; PF01597; GCV_H; 1.
SQ SEQUENCE 30 AA; 3201 MW; 11CE0203120954E6 CRC64;

Query Match 28.7%; Score 29; DB 8; Length 30;
Best Local Similarity 54.5%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 LFYTEEHEIMK 18
|:|:|:|:|:|
Db 7 LAYTSSHEWVK 17

RESULT 8
Q36617 PRELIMINARY; PRT; 30 AA.
AC Q36617:
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Reverse transcriptase (Fragment).
GN POL.
OS Ovine lentivirus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11663;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98042450; PubMed=9375017;
RA Mwaengo D.M., Grant R.F., DeMartini J.C., Carlson J.O.;
RT "Envelope glycoprotein nucleotide sequence and genetic
characterization of North American ovine lentiviruses.";
RL Virology 238:135-144 (1997).
DR EMBL: AF012289; AAB86919.1; -.
FT NON TER 1
FT NON TER 30
SQ SEQUENCE 30 AA; 3650 MW; 94AEDC38E2D80414 CRC64;

Query Match 28.7%; Score 29; DB 15; Length 30;
Best Local Similarity 31.6%; Pred. No. 1.3e+03;
Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKF 19
|:|:|:|:|:|
Db 10 FTMOKILKIEWIEHPMIQF 28

RESULT 9
Q36619 PRELIMINARY; PRT; 30 AA.
AC Q36619:
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Reverse transcriptase (Fragment).
GN POL.
OS Ovine lentivirus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11663;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98042450; PubMed=9375017;
RA Mwaengo D.M., Grant R.F., DeMartini J.C., Carlson J.O.;
RT "Envelope glycoprotein nucleotide sequence and genetic
characterization of North American ovine lentiviruses.";
RL Virology 238:135-144 (1997).
DR EMBL: AF012291; AAB86921.1; -.
FT NON TER 1
FT NON TER 30
SQ SEQUENCE 30 AA; 3664 MW; 8625DC38E39D0414 CRC64;

Query Match 27.7%; Score 28; DB 15; Length 30;
Best Local Similarity 31.6%; Pred. No. 2e+03;
Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKF 19
|:|:|:|:|:|
Db 10 FTMOKILKIEWIEHPMIQF 28

RESULT 10
P82207 PRELIMINARY; PRT; 15 AA.
AC P82207:

```

DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Unknown protein from 2D-page (Fragment).
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
 OC Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=XINHANG X KEMING; TISSUE=Body wall, and Fat body;
 RX MEDLINE=21177481; PubMed=11280994;
 RA Zhong B.X.;
 RT "Protein database for several tissues derived from five instar of
 RT silkworm.";
 RL I Chuan Hsueh Pao 28:217-224 (2001).
 FT NON_TER 15
 SQ SEQUENCE 15 AA; 1877 MW; 580F6BD4703CA70C CRC64;
 Query Match 26.7%; Score 27; DB 5; Length 15;
 Best Local Similarity 44.4%; Pred. No. 1.4e+03;
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 7 SLFYTERHE 15
 |||||:::
 Db 3 SLFYQKQYD 11
 RESULT 11
 OI1394
 ID O14394 PRELIMINARY; PRT; 22 AA.
 AC O14394;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972h-;
 RA Jang Y.-J., Yoo H.-S.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U97394; AAB63886.1; -.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 22 AA; 2510 MW; 48A70026D2982CBA CRC64;
 Query Match 26.7%; Score 27; DB 3; Length 22;
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 EEHEI 16
 |||||
 Db 13 EEHEI 17
 RESULT 12
 O25208
 ID O25208 PRELIMINARY; PRT; 30 AA.
 AC O25208;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein HP0461.
 GN HP0461.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.

OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Peterson S.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Glodek A.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Hickey E.K.,
 RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 RT pylori.";
 RL Nature 388:539-547 (1997).
 DR EMBL; AE000561; AAO07538.1; -.
 DR PIR; E64577; E64577.
 DR TIGR; HP0461; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 30 AA; 3761 MW; 109D0687B1BC8903 CRC64;
 Query Match 26.7%; Score 27; DB 16; Length 30;
 Best Local Similarity 33.3%; Pred. No. 2.9e+03;
 Matches 5; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 1 FSQAKISLFYTEEHE 15
 |||||:::
 Db 15 FKQQLLTFYKTRTHK 29
 RESULT 13
 Q9TRV3
 ID Q9TRV3 PRELIMINARY; PRT; 21 AA.
 AC Q9TRV3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 4-aminobutyrate aminotransferase (Fragment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92041827; PubMed=1939068;
 RA De Biase D., Barra D., Bossa F., Pucci P., John R.A.;
 RT "Chemistry of the inactivation of 4-aminobutyrate aminotransferase by
 RT the antiepileptic drug vigabatrin.";
 RL J. Biol. Chem. 266:20056-20061 (1991).
 DR HSSP; P80147; 1GTX.
 FT NON_TER 1
 SQ SEQUENCE 21 AA; 2500 MW; A0F87150B2DA7A5A CRC64;
 Query Match 26.2%; Score 26.5; DB 6; Length 21;
 Best Local Similarity 31.6%; Pred. No. 2.4e+03;
 Matches 6; Conservative 3; Mismatches 3; Indels 7; Gaps 1;
 QY 9 FYTEEH-----EIMKFS 20
 |||||:::
 Db 1 FWAHEHWGLDDPADVMTFS 19
 RESULT 14
 Q812E2
 ID Q812E2 PRELIMINARY; PRT; 13 AA.
 AC Q812E2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein (Fragment).

GN PFA0655W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Barrinan M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerkhoun A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sultston J.E., Craig A., Newbold C., Barrell B.G;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; AL031744; CAD48947.1; --
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 13 AA; 1619 MW; 50E352E2F7FED1A7 CRC64;

QY 5 KISLFYTEE 13
Db 4 EILTFYSEE 12

Query Match 25.7%; Score 26; DB 5; Length 13;
Best Local Similarity 55.6%; Pred. No. 1.8e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 KISLFYTEE 13
Db 4 EILTFYSEE 12

RESULT 15
Q9RSU5
ID Q9RSU5 PRELIMINARY; PRT; 20 AA.
AC Q9RSU5;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE ACFDLC=FLAVODOXIN (Fragment).
OS Azotobacter chroococcum.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=353;
RN [1]
RP SEQUENCE.
RX MEDLINE=91315397; PubMed=1859358;
RA Bagby S., Barker P.D., Hill H.A., Sanghera G.S., Dunbar B.,
RA Ashby G.A., Eady R.R., Thorneley R.N.;
RT "Direct electrochemistry of two genetically distinct flavodoxins
isolated from Azotobacter chroococcum grown under nitrogen-fixing
conditions.";
RL Biochem. J. 277:313-319(1991).
DR PIR; S17461.
DR InterPro; IPR008254; Flav_nitox_synth.
DR Pfam; PF00258; flavodoxin; 1.
DR PROSITE; PS50902; FLAVODOXIN_LINE; 1.
FT NON TER 1
FT NON TER 20 20
SQ SEQUENCE 20 AA; 2110 MW; 6C8D939F5C999923 CRC64;

Query Match 25.7%; Score 26; DB 2; Length 20;
Best Local Similarity 71.4%; Pred. No. 2.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 AKISLFY 10
||| |

Db 1 AKIGLEFF 7

Search completed: July 15, 2004, 14:50:57
Job time : 27.7333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:47:02 ; Search time 12.4 Seconds
(without alignments)
83.268 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101
Sequence: 1 FSQAKISLFYTERHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 204689

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	32.7	10	1 US-08-240-511-3	Sequence 3, Appli
2	33	32.7	10	2 US-08-884-325-3	Sequence 3, Appli
3	33	32.7	10	3 US-09-290-325-3	Sequence 3, Appli
4	33	32.7	10	4 US-09-382-950-16	Sequence 16, Appli
5	33	32.7	10	4 US-09-382-950-16	Sequence 17, Appli
6	33	32.7	10	4 US-09-382-736B-17	Sequence 3, Appli
7	33	32.7	10	4 US-09-583-243-3	Sequence 3, Appli
8	33	32.7	10	4 US-09-943-120-3	Sequence 3, Appli
9	33	32.7	10	4 US-09-335-018-3	Sequence 3, Appli
10	32	31.7	25	3 US-08-180-371-8	Sequence 8, Appli
11	32	31.7	25	5 PCT-US92-05707-8	Sequence 8, Appli
12	32	31.7	30	3 US-08-180-371-7	Sequence 7, Appli
13	32	31.7	30	5 PCT-US92-05707-7	Sequence 7, Appli
14	31.5	31.2	19	1 US-08-097-997A-5	Sequence 5, Appli
15	31.5	31.2	19	3 US-08-665-574C-5	Sequence 5, Appli
16	31.5	31.2	19	3 US-08-946-994-5	Sequence 5, Appli
17	31.5	31.2	20	2 US-08-683-743-22	Sequence 22, Appli
18	30	29.7	13	2 US-08-572-951-4	Sequence 4, Appli
19	30	29.7	13	2 US-08-572-951-21	Sequence 21, Appli
20	30	29.7	20	3 US-08-861-338-7	Sequence 7, Appli
21	30	29.7	21	1 US-08-178-606-6	Sequence 6, Appli
22	29	28.7	9	1 US-08-454-207A-32	Sequence 32, Appli
23	28	27.7	10	1 US-08-129-610-9	Sequence 9, Appli
24	28	27.7	10	1 US-08-129-609A-9	Sequence 9, Appli
25	28	27.7	10	1 US-08-455-313-9	Sequence 9, Appli
26	28	27.7	10	3 US-09-224-025-9	Sequence 9, Appli
27	28	27.7	10	5 PCT-US94-07887-9	Sequence 9, Appli

28 28 27.7 12 3 US-08-742-243-7 Sequence 7, Appli
29 28 27.7 12 3 US-08-742-243-8 Sequence 8, Appli
30 28 27.7 12 3 US-08-742-243-9 Sequence 9, Appli
31 28 27.7 13 2 US-08-572-951-22 Sequence 22, Appli
32 28 27.7 15 2 US-08-480-190-228 Sequence 228, App
33 28 27.7 15 2 US-08-488-379-228 Sequence 228, App
34 28 27.7 15 4 US-08-475-399A-228 Sequence 228, App
35 28 27.7 15 5 PCT-US93-07545-228 Sequence 228, App
36 28 27.7 17 2 US-08-719-758-7 Sequence 7, Appli
37 28 27.7 17 3 US-09-119-827-7 Sequence 7, Appli
38 28 27.7 20 1 US-08-218-026-11 Sequence 11, Appli
39 28 27.7 20 2 US-08-653-632-11 Sequence 11, Appli
40 28 27.7 21 1 US-08-218-026-12 Sequence 12, Appli
41 28 27.7 21 2 US-08-653-632-12 Sequence 12, Appli
42 28 27.7 21 6 5194596-11 Patent No. 5194596
43 28 27.7 21 6 5219739-11 Patent No. 5219739
44 28 27.7 22 1 US-08-649-272A-6 Sequence 6, Appli
45 28 27.7 22 1 US-08-218-026-59 Sequence 59, Appli

ALIGNMENTS

RESULT 1
US-08-240-511-3
; Sequence 3, Application US/08240511
; Patent No. 5643722
; GENERAL INFORMATION:
; APPLICANT: ROTHSCCHILD, Kenneth J.
; APPLICANT: SONAR, Sanjay M.
; TITLE OF INVENTION: METHODS FOR THE DETECTION AND ISOLATION OF PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,511
; FILING DATE: 11-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 16865-0128
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 10 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-240-511-3

Query Match 32.7%; Score 33; DB 1; Length 10;

Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 YTEHEIMK 18
Db ||::||
2 YTKDHDIRK 10

RESULT 2

US-08-884-325-3
; Sequence 3, Application US/08884325
; Patent No. 5922858
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth J.
; APPLICANT: Sonar, Sanjay M.
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods for the Detection and Isolation
; TITLE OF INVENTION: of Proteins
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,325
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/240,511
; FILING DATE: 11-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: AMBER-03242
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-705-8410
; TELEFAX: 415-397-8338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "The amino acid in this
; OTHER INFORMATION: position is N-formylmethionine"
US-08-884-325-3

Query Match 32.7%; Score 33; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 YTEHEIMK 18
Db ||::||
2 YTKDHDIRK 10

RESULT 3

US-09-290-325-3
; Sequence 3, Application US/09290325
; Patent No. 6210941
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth
; APPLICANT: Sonar, Sanjay M.
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods for the Detection and Isolation
; TITLE OF INVENTION: of Proteins
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/290,325
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/240,511
; FILING DATE: 11-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: AMBER-03242
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-705-8410
; TELEFAX: 415-397-8338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "The amino acid in this
; OTHER INFORMATION: position is N-formylmethionine"
US-09-290-325-3

Query Match 32.7%; Score 33; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 YTEHEIMK 18
Db ||::||
2 YTKDHDIRK 10

; APPLICANT: Rothschild, Kenneth J.
; APPLICANT: Sonar, Sanjay M.
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods for the Detection and Isolation
; TITLE OF INVENTION: of Proteins
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/290,325
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,325
; FILING DATE: 27-JUN-1997
; APPLICATION NUMBER: US 08/240,511
; FILING DATE: 11-MAY-1994
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: AMBER-03242
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-705-8410
; TELEFAX: 415-397-8338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 6210941 Relevant
; TOPOLOGY: No. 6210941 Relevant
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "The amino acid in this
; OTHER INFORMATION: position is N-formylmethionine"
US-09-290-325-3

Query Match 32.7%; Score 33; DB 3; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 YTEHEIMK 18
Db ||::||
2 YTKDHDIRK 10

RESULT 4

US-09-382-950-16
; Sequence 16, Application US/09382950
; Patent No. 6303337
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth
; APPLICANT: Gite, Sadanand
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: N-Terminal and C-Terminal Markers in Nascent Proteins
; FILE REFERENCE: AMBER-03879
; CURRENT APPLICATION NUMBER: US/09/382,950
; CURRENT FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 10

Query Match 32.7%; Score 33; DB 3; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 YTEHEIMK 18
Db ||::||
2 YTKDHDIRK 10

```
;
; TYPE: PRT
; ORGANISM: Artificial/Unknown
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: ()
; OTHER INFORMATION: Synthetic
US-09-382-950-16

Query Match          32.7%; Score 33; DB 4; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      10 YTEEHEIMK 18
Db      2 YTKDHDIRK 10

RESULT 5
US-09-382-736B-17
; Sequence 17, Application US/09382736B
; Patent No. 6306628
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth
; APPLICANT: Gite, Sadanand
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods for the Detection, Analysis and Isolation of Nascent Proteins
; FILE REFERENCE: AMBER-03951
; CURRENT APPLICATION NUMBER: US/09/382,736B
; CURRENT FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 17
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-382-736B-17

Query Match          32.7%; Score 33; DB 4; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      10 YTEEHEIMK 18
Db      2 YTKDHDIRK 10

RESULT 6
US-09-307-579-3
; Sequence 3, Application US/09307579
; Patent No. 6344320
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth J.
; APPLICANT: Sonar, Sanjay M.
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods for the Detection and Isolation
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/307,579
```

```
;
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/884,325
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: AMBER-03242
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-705-8410
; TELEFAX: 415-397-8338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "The amino acid in this
; position is N-formylmethionine"
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-307-579-3

Query Match          32.7%; Score 33; DB 4; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      10 YTEEHEIMK 18
Db      2 YTKDHDIRK 10

RESULT 7
US-09-583-243-3
; Sequence 3, Application US/09583243
; Patent No. 6358689
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Detection of Marker in Nascent Proteins
; FILE REFERENCE: AMBER-04437
; CURRENT APPLICATION NUMBER: US/09/583,243
; CURRENT FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/335,018
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-583-243-3

Query Match          32.7%; Score 33; DB 4; Length 10;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      10 YTEEHEIMK 18
Db      2 YTKDHDIRK 10

RESULT 8
US-09-943-120-3
; Sequence 3, Application US/09943120
; Patent No. 6566070
; GENERAL INFORMATION:
```

APPLICANT: Rothschild, Kenneth J.
 Sonar, Sanjay M.
 Olejnik, Jerzy
 TITLE OF INVENTION: Methods for the Detection and Isolation
 of Proteins
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MEDLEN & CARROLL, LLP
 STREET: 220 Montgomery Street, Suite 2200
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/943,120
 FILING DATE: 30-Aug-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/307,579
 FILING DATE: 07-May-1999
 APPLICATION NUMBER: 08/884,325
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Carroll, Peter G.
 REGISTRATION NUMBER: 32,837
 REFERENCE/DOCKET NUMBER: AMBER-03242
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-705-8410
 TELEFAX: 415-397-8338
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Modified-site
 LOCATION: 1
 OTHER INFORMATION: /note="The amino acid in this
 position is N-formylmethionine"
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-09-943-120-3
 Query Match 32.7%; Score 33; DB 4; Length 10;
 Best Local Similarity 55.6%; Pred. No. 18;
 Matches 5; Conservative 3; Mismatches 1; Indels 0;
 QY 10 YTEHEIMK 18
 ||:|:|
 DB 2 YTKDHDIRK 10
 RESULT 9
 US/09-335-018-3
 ; Sequence 3, Application US/09335018
 ; Patent No. 6596481
 ; GENERAL INFORMATION:
 ; APPLICANT: Rothschild, Kenneth J.
 ; Sonar, Sanjay M.
 ; Olejnik, Jerzy
 ; TITLE OF INVENTION: Methods for the Detection and Isolation
 ; of Proteins
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MEDLEN & CARROLL, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 906866
; FILING DATE: 01 July 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gruber, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27620/31668
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-180-371-8
Query Match 31.7%; Score 32; DB 3; Length 25;
Best Local Similarity 42.9%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 KISLFYTEHEIMK 18
Db 5 KINLFYNSNPWK 18

RESULT 11
PCT-US92-05707-8
; Sequence 8, Application PC/TUS9205707
; GENERAL INFORMATION:
; APPLICANT: Choudhury, Chandra
; TITLE OF INVENTION: Hematopoietic Growth Factor Derived from
; TITLE OF INVENTION: T Lymphocytes and Methods of Use Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05707
; FILING DATE: 19920707
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gruber, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27620/30933
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-05707-8
Query Match 31.7%; Score 32; DB 5; Length 25;
Best Local Similarity 42.9%; Pred. No. 73;

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 906866
; FILING DATE: 01 July 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gruber, Lewis S.
; REGISTRATION NUMBER: 30,060
; REFERENCE/DOCKET NUMBER: 27620/31668
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-180-371-7
Query Match 31.7%; Score 32; DB 3; Length 30;
Best Local Similarity 42.9%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 KISLFYTEHEIMK 18
Db 7 KINLFYNSNPWK 20

RESULT 13
PCT-US92-05707-7
; Sequence 7, Application PC/TUS9205707
; GENERAL INFORMATION:
; APPLICANT: Choudhury, Chandra
; TITLE OF INVENTION: Hematopoietic Growth Factor Derived from
; TITLE OF INVENTION: T Lymphocytes and Methods of Use Therefor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark

```

STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/05707
FILING DATE: 19920707
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gruber, Lewis S.
REGISTRATION NUMBER: 30,060
REFERENCE/DOCKET NUMBER: 27620/30933
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-05707-7

Query Match 31.7%; Score 32; DB 5; Length 30;
Best Local Similarity 42.9%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY

5 KISLFYTEEHEIMK 18
|||||
7 KINLFYNSPVMK 20

Db

RESULT 14

US-08-097-997A-5
Sequence 5, Application US/08097997A
Patent No. 5728536
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvennoinen, Ollie
APPLICANT: Witthuhn, Bruce A.
APPLICANT: Quelle, Frederick W.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
TITLE OF INVENTION: Transduction
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,997A
FILING DATE: 29-JULY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0370000/SLF/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-097-997A-5

Query Match 31.2%; Score 31.5; DB 1; Length 19;
Best Local Similarity 40.0%; Pred. No. 66;
Matches 6; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

QY 2 SQAKISLFYTEEHEI 16
||| : || : || :
Db 2 SQRKIQ-FYEDKHQL 15

RESULT 15

US-08-665-574C-5
Sequence 5, Application US/08665574C
Patent No. 6136595
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvennoinen, Ollie
APPLICANT: Witthuhn, Bruce A.
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine
TITLE OF INVENTION: Signal Transduction
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,574C
FILING DATE: 18-JUN-1996
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/282,012
FILING DATE: 29-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/097,997
FILING DATE: 29-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/118,968
FILING DATE: 09-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/LBB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-665-574C-5

Query Match 31.2%; Score 31.5; DB 3; Length 19;
Best Local Similarity 40.0%; Pred. No. 66;
Matches 6; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

QY 2 SQAKISLFYTEEHEI 16
||| : || : || :
Db 2 SQRKIQ-FYEDKHQL 15

Db 2 SQKLO-FYEDKHQL 15

Search completed: July 15, 2004, 15:13:20
Job time : 13.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:11:48 ; Search time 32.5333 Seconds

(without alignments)
192.148 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101

Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 288454

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	10	US-09-171-432A-43
2	42	41.6	20	10	US-09-171-432A-42
3	34	33.7	25	16	US-10-416-249-401
4	33	32.7	10	9	US-09-973-145-16
5	33	32.7	10	14	US-10-174-388A-16
6	33	32.7	10	14	US-10-264-127-17
7	33	32.7	10	15	US-10-345-664-6
8	33	32.7	10	15	US-10-339-712-83
9	32	31.7	28	14	US-10-092-750-60
10	31	30.7	18	14	US-10-225-567A-2273
11	31	30.7	25	16	US-10-416-249-101
12	31	30.7	25	16	US-10-416-249-332
13	30	29.7	16	12	US-10-346-162-201
14	30	29.7	18	14	US-10-084-813-847
15	30	29.7	20	9	US-09-736-076-7

15	30	29.7	21	14	US-10-084-813-585	Sequence 585, App
17	30	29.7	21	14	US-10-084-813-586	Sequence 586, App
18	30	29.7	21	14	US-10-084-813-559	Sequence 559, App
19	30	29.7	21	14	US-10-084-813-560	Sequence 560, App
20	30	29.7	21	14	US-10-084-813-561	Sequence 561, App
21	30	29.7	21	14	US-10-084-813-562	Sequence 562, App
22	30	29.7	28	12	US-10-424-599-158169	Sequence 158169,
23	30	29.7	30	9	US-09-864-761-36666	Sequence 36666, A
24	29.5	29.2	25	14	US-10-029-386-33285	Sequence 33285, A
25	29.5	29.2	25	16	US-10-416-249-403	Sequence 403, App
26	29	28.7	15	14	US-10-084-813-801	Sequence 801, App
27	29	28.7	15	14	US-10-084-813-802	Sequence 802, App
28	29	28.7	15	16	US-10-644-703-16	Sequence 16, Appl
29	29	28.7	17	16	US-10-644-703-17	Sequence 17, Appl
30	29	28.7	18	10	US-09-962-863-11	Sequence 11, Appl
31	29	28.7	18	14	US-10-084-813-848	Sequence 848, App
32	29	28.7	18	14	US-10-084-813-849	Sequence 849, App
33	29	28.7	18	14	US-10-084-813-850	Sequence 850, App
34	29	28.7	18	14	US-10-084-813-851	Sequence 851, App
35	29	28.7	21	11	US-09-754-831A-19	Sequence 19, Appl
36	29	28.7	21	14	US-10-084-813-324	Sequence 324, App
37	29	28.7	21	14	US-10-084-813-587	Sequence 587, App
38	29	28.7	21	14	US-10-084-813-588	Sequence 588, App
39	29	28.7	21	14	US-10-084-813-589	Sequence 589, App
40	29	28.7	21	14	US-10-084-813-590	Sequence 590, App
41	29	28.7	21	14	US-10-084-813-663	Sequence 663, App
42	29	28.7	21	14	US-10-084-813-664	Sequence 664, App
43	29	28.7	21	14	US-10-084-813-665	Sequence 665, App
44	29	28.7	21	14	US-10-084-813-666	Sequence 666, App
45	29	28.7	25	15	US-10-360-101-284	Sequence 284, App

ALIGNMENTS

RESULT 1

US-09-171-432A-43
; Sequence 43, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1-20
OTHER INFORMATION: /label= YK-1319
US-09-171-432A-43

Query Match 100.0%; Score 101; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHIMKFS 20
DB 1 FSOAKISLFYTEEHIMKFS 20

RESULT 2

US-09-171-432A-42
; Sequence 42, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171.432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1-20
; OTHER INFORMATION: /label= YK-1318
US-09-171-432A-42

Query Match 41.6%; Score 42; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLF 9
DB 12 FSOAKISLF 20

RESULT 3
US-10-416-249-401
; Sequence 401, Application US/10416249
; Publication No. US20040132121A1
; GENERAL INFORMATION:
; APPLICANT: Dalrymple, Brian P.
; APPLICANT: Kongsuwan, Kritaya
; APPLICANT: Wijffels, Gene L.
; APPLICANT: Jennings, Philip A.
; APPLICANT: Kemp, Gregory W.
; TITLE OF INVENTION: METHOD OF IDENTIFYING ANTIBACTERIAL
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: CULLM42.001APC
; CURRENT APPLICATION NUMBER: US/10/416,249
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: PCT/AU01/01436
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: AU PR 1320
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: AU PR 2919
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 678
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 401
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-416-249-401

Query Match 33.7%; Score 34; DB 16; Length 25;
Best Local Similarity 56.2%; Pred. No. 1.6e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 3 QAKISLFYTEE--HEI 16
DB 9 QGQSLFGDEEKAHEI 24

RESULT 4

US-09-973-145-16
; Sequence 16, Application US/09973145
; Patent No. US20020132248A1
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth J.
; APPLICANT: Gite, Sadanand
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: N-Terminal and C-Terminal Markers in Nascent Proteins
; FILE REFERENCE: AMBER-08819
; CURRENT APPLICATION NUMBER: US/09/973,145
; CURRENT FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 09/382,950
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-973-145-16

Query Match 32.7%; Score 33; DB 9; Length 10;
Best Local Similarity 55.6%; Pred. No. 84;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 YTEEHIMK 18
|:::|

Db 2 YTKDHDIRK 10

RESULT 5

US-10-174-368A-16
; Sequence 16, Application US/10174368A
; Publication No. US20030092031A1
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth J.
; APPLICANT: Gite, Sadanand
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods for the Detection, Analysis and Isolation of Nascent Proteins
; FILE REFERENCE: AMBER-07145
; CURRENT APPLICATION NUMBER: US/10/174,368A
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 10/049,332
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/23233
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/382,736
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-174-368A-16

Query Match 32.7%; Score 33; DB 14; Length 10;
Best Local Similarity 55.6%; Pred. No. 84;
Matches 5; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

Qy 10 YTEBHEIMK 18
||:|:|
Db 2 YTKDHDIRK 10

RESULT 6

US-10-264-127-17
; Sequence 17, Application US/10264127
; Publication No. US20030190643A1
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth
; APPLICANT: Gite, Sadanand
; APPLICANT: Olejnik, Jerzy
; TITLE OF INVENTION: Methods for the Detection, Analysis and Isolation of Nascent Proteins
; FILE REFERENCE: AMBER-03951
; CURRENT APPLICATION NUMBER: US/10/264,127
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: US/09/382,736B
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-264-127-17

Query Match 32.7%; Score 33; DB 14; Length 10;
Best Local Similarity 55.6%; Pred. No. 84;
Matches 5; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

Qy 10 YTEBHEIMK 18
||:|:|
Db 2 YTKDHDIRK 10

RESULT 7

US-10-345-664-6
; Sequence 6, Application US/10345664
; Publication No. US20030219780A1
; GENERAL INFORMATION:
; APPLICANT: Olejnik, Jerzy
; APPLICANT: Krzymanska-Olejnik, Edyta
; APPLICANT: Mamaev, Sergey
; APPLICANT: Rothschild, Kenneth J.
; TITLE OF INVENTION: Methods for the Preparation of Chemically Mismatched tRNA
; FILE REFERENCE: AMBER-07733
; CURRENT APPLICATION NUMBER: US/10/345,664
; CURRENT FILING DATE: 2003-01-16
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-345-664-6

Query Match 32.7%; Score 33; DB 15; Length 10;
Best Local Similarity 55.6%; Pred. No. 84;
Matches 5; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

Qy 10 YTEBHEIMK 18
||:|:|
Db 2 YTKDHDIRK 10

RESULT 8

US-10-339-712-83
; Sequence 83, Application US/10339712
; Publication No. US20040014071A1
; GENERAL INFORMATION:
; APPLICANT: Rothschild, Kenneth J.
; APPLICANT: Gite, Sadanand
; APPLICANT: Olejnik, Jerzy
; APPLICANT: Lim, Mark
; TITLE OF INVENTION: Methods for the Detection, Analysis and Isolation of Nascent Proteins
; FILE REFERENCE: AMBER-07199
; CURRENT APPLICATION NUMBER: US/10/339,712
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/049,322
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: PCT/US00/23233
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/382,736
; PRIOR FILING DATE: 1999-08-25
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 83
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-339-712-83

Query Match 32.7%; Score 33; DB 15; Length 10;
Best Local Similarity 55.6%; Pred. No. 84;
Matches 5; Conservative 3; Mismatches 1; Indels 1; Gaps 0;

Qy 10 YTEBHEIMK 18
||:|:|
Db 2 YTKDHDIRK 10

RESULT 9

```
US-10-092-750-60
; Sequence 60, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-X1
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-60

Query Match      31.7%; Score 32; DB 14; Length 28;
Best Local Similarity 45.5%; Pred. No. 3.8e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      5 KISLFYTEEHE 15
Db      16 KLIKFYSKKE 26

RESULT 10
US-10-225-567A-2273
; Sequence 2273, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenn A.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2273
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-2273

Query Match      30.7%; Score 31; DB 14; Length 18;
Best Local Similarity 45.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      6 ISLFYTEEHEI 16
Db      2 VSLAYSRSHQI 12

RESULT 11
US-10-416-249-101
; Sequence 101, Application US/10416249
; Publication No. US20040132121A1
; GENERAL INFORMATION:
; APPLICANT: Dalrymple, Brian P.
; APPLICANT: Kongsuwan, Kritaya
; APPLICANT: Wijffels, Gene L.
; APPLICANT: Jennings, Philip A.
; APPLICANT: Kemp, Gregory W.
; TITLE OF INVENTION: METHOD OF IDENTIFYING ANTIBACTERIAL
; FILE REFERENCE: CULLN42.001APC
; CURRENT APPLICATION NUMBER: US/10/416,249
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: PCT/AU01/01436
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: AU PR 1320
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: AU PR 2919
; NUMBER OF SEQ ID NOS: 678
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Chlorobium tepidum
US-10-416-249-332

Query Match      30.7%; Score 31; DB 16; Length 25;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy      5 KISLFYTEEHEIMK 18
Db      11 QISLFEEESRLRK 24

RESULT 13
US-10-346-162-201
; Sequence 201, Application US/10346162
; Publication No. US20030224390A1
; GENERAL INFORMATION:
; APPLICANT: KARO BIO USA, INC.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: BARNETT, Thomas R.
; TITLE OF INVENTION: METHOD OF IDENTIFYING ANTIBACTERIAL
```

; APPLICANT: BUEHRER, Benjamin
; TITLE OF INVENTION: METHOD OF IDENTIFYING CONFORMATION-SENSITIVE BINDING PEPTIDES AND
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: PAIGE=1H
; CURRENT APPLICATION NUMBER: US/10/346,162
; CURRENT FILING DATE: 2003-01-17
; PRIOR APPLICATION NUMBER: US 09/614,865
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 09/860,688
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 201
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-346-162-201

Query Match 29.7%; Score 30; DB 12; Length 16;
Best Local Similarity 28.6%; Pred. No. 4.4e+02;
Matches 4; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 7 SLFYTEEHEIMKFS 20
| : : : | : : : |
Db 1 SRWLDDEHLLYS 14

RESULT 14
US-10-084-813-847
; Sequence 847, Application US/10084813
; Publication No. US2003006815A1
; GENERAL INFORMATION:
; APPLICANT: SAXINGER, CARL
; TITLE OF INVENTION: POLYPEPTIDES THAT BIND HIV Gp120 AND RELATED NUCLEIC
; TITLE OF INVENTION: ACIDS, ANTIBODIES, COMPOSITIONS, AND METHODS OF USE
; FILE REFERENCE: 215875
; CURRENT APPLICATION NUMBER: US/10/084,813
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/US00/23505
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,270
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 1242
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 847
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: binding peptide
US-10-084-813-847

Query Match 29.7%; Score 30; DB 14; Length 18;
Best Local Similarity 38.9%; Pred. No. 5e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 QAKISLFYTEEHEIMKFS 20
| : : : | : : : |
Db 1 QKASSIVYKGEQVEFS 18

RESULT 15
US-09-736-076-7
; Sequence 7, Application US/09736076
; Patent No. US20020049301A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson Shmuel A.
; TITLE OF INVENTION: SHORT PEPTIDES WHICH SELECTIVELY
; TITLE OF INVENTION: MODULATE THE ACTIVITY OF SERINE/THREONINE KINASES
; FILE REFERENCE: 1242.1015-009
; CURRENT APPLICATION NUMBER: US/09/736,076

; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: US 08/861,338
; PRIOR FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Akt/PKB
US-09-736-076-7

Query Match 29.7%; Score 30; DB 9; Length 20;
Best Local Similarity 57.1%; Pred. No. 5.6e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 FYTEEHE 15
| : : : | : : : |
Db 10 FYNQDHE 16

Search completed: July 15, 2004, 15:22:55
Job time : 33.5333 secs

10/10/2010 10:10:10

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:36:31 ; Search time 40.1333 Seconds
(without alignments)
140.804 Million cell updates/sec

Title: US-09-171-432A-44
Perfect score: 104
Sequence: 1 KVNFPHGMLDLEITAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 581562

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	20	2	AAW42927 Immunogen
2	104	100.0	21	4	ABP69444 Synthetic
3	53	51.0	20	2	AAW42928 Immunogen
4	53	51.0	21	4	ABP69445 Synthetic
5	33	31.7	14	4	AAW98649 Human pep
6	33	31.7	14	5	ABP46502 Human Bly
7	33	31.7	16	6	ABP82445 G protein
8	33	31.7	22	3	ABP51604 Yada homo
9	33	31.7	24	3	ABP51532 Proteobac
10	32	30.8	8	2	AAW42927 Immunogen
11	32	30.8	9	3	AAW42928 Immunogen
12	32	30.8	10	3	AAW42928 Immunogen
13	32	30.8	10	3	AAW42928 Immunogen
14	32	30.8	14	5	ABP46141 Human Bly
15	32	30.8	19	5	AAO21225 Alpha-iso
16	31	29.8	23	7	ADT73708 Yeast Gic
17	31	29.8	20	2	AAW39834 E1 peptid
18	31	29.8	20	4	AAW74089 C-termina
19	31	29.8	21	2	AAW65690 Fibronect
20	31	29.8	22	1	AAW65676 Fibronect
21	31	29.8	24	2	AAW05554 N-terminu
22	31	29.8	30	5	AAU85095 Human PRA
23	31	29.8	30	5	AAU85094 Human PRA
24	30.5	29.3	22	2	AAW38482 Human sec
25	30	28.8	12	7	ADC28091 Synthetic

ALIGNMENTS

RESULT 1
AAW42927
ID AAW42927 standard; peptide; 20 AA.
XX
AC AAW42927;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1327.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006891.
XX
PR 19-APR-1996; 96US-0015644P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Fields HA, Khudyakov YE;
XX
XX WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune response to HAV in a mammal or to detect the presence of antibodies against HAV in a mammal.
XX
XX Claim 18; Page 112; 140pp; English.
XX
XX Peptides AAW42922-30 are immunogenic peptides corresponding to immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are substantially similar to a portion of the amino acid sequence of the P2A protein of HAV corresponding to amino acids 732-980. The present peptide is derived from amino acids 922-941, and has a reactivity of 31.3% with acute sera. Compositions containing the peptides can be used to induce an immune response to HAV in a mammal. The peptides can also be used to detect the presence of antibodies against HAV in mammalian serum. The peptides can also be used to make an antibody against HAV by administering the peptide to a mammal

Sequence 20 AA;

26 30 28.8 14 4 AAW96766 Human pep
27 30 28.8 14 5 ABP46462 Human Bly
28 30 28.8 14 5 AAU93647 Granulocy
29 30 28.8 14 6 ABP72195 Peptide a
30 30 28.8 15 5 ABP55175 Human mit
31 30 28.8 17 2 AAR05760 Peptide a
32 30 28.8 19 2 AAW70492 Aspergill
33 30 28.8 26 7 ADA07512 Human sec
34 29.5 28.4 17 7 ADB98809 LRP5 pept
35 29 27.9 7 4 AAE09073 Phage pep
36 29 27.9 13 4 AAU68418 Human Bre
37 29 27.9 13 5 ABG67808 Human ADP
38 29 27.9 13 6 ADA23948 Alzheimer
39 29 27.9 14 4 AAM99297 Vaccine r
40 29 27.9 14 5 ABP46527 Human Bly
41 29 27.9 14 5 ABP46391 Human Bly
42 29 27.9 14 5 ABP46464 Human Bly
43 29 27.9 14 5 ABP46393 Human Bly
44 29 27.9 14 5 ABP46358 Human Bly
45 29 27.9 14 5 ABP46389 Human Bly

Query Match 100.0%; Score 104; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEIIAANSKD 20
 |||||
 Db 1 KVNFPFGMLDLEIIAANSKD 20

RESULT 2
 AAB69444
 ID AAB69444 standard; peptide; 21 AA.
 XX AC AAB69444;
 XX XX
 DT 20-APR-2001 (first entry)
 XX XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 44.
 XX XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 XX antigen; major structural capsid polypeptide; HAV antibody detection.
 XX OS Hepatitis A virus.
 XX OS Synthetic.
 XX PN WO200105824-A2.
 XX XX
 PD 25-JAN-2001.
 XX XX
 PF 14-JUL-2000; 2000WO-US019267.
 XX XX
 PR 15-JUL-1999; 99US-0144412P.
 XX XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PA Fields HA, Khudyakov YE;
 XX PI
 XX DR WFI; 2001-112681/12.
 XX XX
 PT Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 XX XX
 PS Claim 13; Page 96; 130pp; English.
 XX XX
 CC The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IgM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IgM antibody reactivity

Sequence 21 AA;

Query Match 100.0%; Score 104; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.3e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEIIAANSKD 20
 |||||
 Db 1 KVNFPFGMLDLEIIAANSKD 20

RESULT 3
 AAW42928
 ID AAW42928 standard; peptide; 20 AA.
 XX AC AAW42928;
 XX XX
 DT 28-APR-1998 (first entry)
 XX XX
 DE Immunogenic Hepatitis A virus peptide YK-1328.
 XX XX
 KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
 KW antibody.
 XX OS Synthetic.
 XX OS Hepatitis A virus.
 XX PN WO9740147-A1.
 XX XX
 PD 30-OCT-1997.
 XX XX
 PF 18-APR-1997; 97WO-US006891.
 XX XX
 PR 19-APR-1996; 96US-0015644P.
 XX XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PA Fields HA, Khudyakov YE;
 XX PI
 XX DR WFI; 1997-535831/49.
 XX XX
 PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.
 XX XX
 PS Claim 18; Page 112; 140pp; English.
 XX XX
 CC Peptides AAW42922-30 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the P2A
 CC protein of HAV corresponding to amino acids 792-980. The present peptide
 CC is derived from amino acids 931-950, and has a reactivity of 12.5% with
 CC acute sera. Compositions containing the peptides can be used to induce an
 CC immune response to HAV in a mammal. The peptides can also be used to
 CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal

Sequence 20 AA;

Query Match 51.0%; Score 53; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.088;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 DLEBIAANSKD 20
 |||||
 Db 1 DLEBIAANSKD 11

RESULT 4
 AAB69445
 ID AAB69445 standard; peptide; 21 AA.
 XX AC AAB69445;
 XX XX
 DT 20-APR-2001 (first entry)
 XX XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 45.
 XX XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 XX antigen; major structural capsid polypeptide; HAV antibody detection.
 XX OS Hepatitis A virus.
 XX OS Synthetic.

XX WO200105824-A2.
 PN 25-JAN-2001.
 XX
 PD 14-JUL-2000; 2000WO-US019267.
 PF 15-JUL-1999; 99US-0144412P.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Fields HA, Khudyakov YE;
 PI WPI; 2001-112681/12.
 XX Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 PT Claim 13; Page 97; 130pp; English.
 PS
 XX The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IGM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IGM antibody reactivity
 XX
 SQ Sequence 21 AA;
 Query Match 51.0%; Score 53; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.093;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 DLEETAAASKD 20
 Db |||||
 1 DLEETAAASKD 11
 RESULT 5
 AAM98649
 ID AAM98649 standard; peptide; 14 AA.
 XX
 AC AAM98649;
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE Human peptide #1924 encoded by a SNP oligonucleotide.
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease.
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 PD 05-JUL-2001.
 XX

PF 28-DEC-2000; 2000WO-US035498.
 XX
 PR 28-DEC-1999; 99US-0173419P.
 PR 27-DEC-2000; 2000US-00173419.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2001-465210/50.
 XX
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
 PT autoimmune diseases and infections.
 XX
 PS Disclosure; Page 4090; 4143pp; English.
 XX
 CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, cytochromes, kinesins, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms
 XX
 SQ Sequence 14 AA;
 Query Match 31.7%; Score 33; DB 4; Length 14;
 Best Local Similarity 62.5%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 VNFPHGML 9
 Db : |||||
 4 IQFPHGRL 11
 RESULT 6
 ABP46502
 ID ABP46502 standard; peptide; 14 AA.
 XX
 AC ABP46502;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human BlyS binding scFv VH CDR3 SEQ ID 2513.
 XX
 KW BlyS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 XX common variable immunodeficiency; acquired immunodeficiency syndrome.
 OS Homo sapiens.
 XX
 PN WO200202641-A1.
 XX
 PD 10-JAN-2002.
 XX
 PF 15-JUN-2001; 2001WO-US019110.
 XX
 PR 16-JUN-2000; 2000US-0212210P.
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR

```

PR 21-MAR-2001; 2001US-0277379P.
PR 25-MAY-2001; 2001US-0293499P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
PI WPI; 2002-114799/15.
XX
XX Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
PT diagnosis and treatment of cancers and immune disorders.
PT
XX Claim 2; Page 3003; 3148pp; English.
PS
XX This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (BLyS) polypeptides. BLyS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BLyS. The antibodies bind to BLyS
CC and so may be used to detect and quantitate the presence of BLyS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BLyS. They may also be
CC administered to treat diseases associated with aberrant BLyS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method of
CC the invention
XX
XX Sequence 14 AA;
SQ
Query Match 31.7%; Score 33; DB 5; Length 14;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 FPHGMDDL 11
DB ||| |||
7 FPHAPLGL 14
RESULT 7
ABP82445
ID ABP82445 standard; peptide; 16 AA.
XX AC
XX ABP82445;
XX 04-MAR-2003 (first entry)
XX
XX G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1118.
XX
XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer.
XX
XX Homo sapiens.
XX OS
XX WO200261087-A2.
XX PN
XX 08-AUG-2002.
XX PD
XX 19-DEC-2001; 2001WO-US050107.
XX PF

```

```

XX 19-DEC-2000; 2000US-0257144P.
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX Burmer GC, Roush CL, Brown JP;
XX WPI; 2003-046718/04.
XX
XX New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.
XX
XX Claim 1; Fig 2; 523pp; English.
XX
XX The present invention describes antigenic peptides (I) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
XX Sequence 16 AA;
SQ
Query Match 31.7%; Score 33; DB 6; Length 16;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 5 PHGMDDL 13
DB ||| |||
5 PHSVIDVEE 13
RESULT 8
AAB51604
ID AAB51604 standard; peptide; 22 AA.
XX AC
XX AAB51604;
XX 15-FEB-2001 (first entry)
XX YadaA homologous peptide #67.
XX Proteobacteria; extracellular domain; virulence determinant; Yada;
KW adhesin; proteobacterial infection prevention; vaccine.
XX
XX Actinobacillus actinomycetemcomitans.
XX OS
XX WO200061165-A1.
XX PN
XX 19-OCT-2000.
XX PD
XX 13-APR-2000; 2000WO-US009966.
XX PF

```

PR 13-APR-1999; 99US-0129073P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX Lupas AN;
 PI WPI; 2000-647397/62.
 DR An isolated polypeptide conserved in proteobacterial extracellular
 PT domains used in the treatment and prevention of bacterial infections.
 PT Example 5; Page 60; 85pp; English.
 XX This invention relates to peptides AAB51512 - AAB51537 which represent
 CC conserved proteobacterial extracellular domains. Sequences AAB51538 -
 CC AAB51618 represent peptides homologous to Yada, a Yersinia adhesin which
 CC is an important virulence determinant of the Yersinia species. The
 CC invention includes an antibody which binds to the proteobacterial
 CC extracellular peptides, and an immunogenic composition containing the
 CC antibody used as a vaccine to prevent infection by a proteobacteria. The
 CC polypeptides and antibodies are useful in the treatment and prevention of
 CC proteobacterial infections. The polypeptides can also be used to identify
 CC compounds which antagonize the binding of a bacterial adhesion to its
 CC ligand. The host cell can be used to produce the polypeptides in a
 CC suitable culture system. The composition can be used to vaccinate a
 CC patient against a proteobacterial infection
 XX Sequence 22 AA;
 SQ

Query Match 31.7%; Score 33; DB 3; Length 22;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 TAANSKD 20
 |||||
 Db 9 TAANSKD 15

RESULT 9
 AAB51532
 ID AAB51532 standard; peptide; 24 AA.
 AC AAB51532;
 XX 15-FEB-2001 (first entry)
 DT Proteobacterial extracellular conserved peptide SEQ ID 21.
 DE Proteobacteria; extracellular domain; virulence determinant; Yada;
 KW adhesin; proteobacterial infection prevention; vaccine.
 XX Actinobacillus actinomycetemcomitans.
 OS WO200061165-A1.
 XX 19-OCT-2000.
 PD 13-APR-2000; 2000WO-US009866.
 PF 13-APR-1999; 99US-0129073P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX Lupas AN;
 PI WPI; 2000-647397/62.
 DR An isolated polypeptide conserved in proteobacterial extracellular
 PT domains used in the treatment and prevention of bacterial infections.
 XX

PS Claim 8; Page 65; 85pp; English.
 XX This invention relates to peptides AAB51512 - AAB51537 which represent
 CC conserved proteobacterial extracellular domains. Sequences AAB51538 -
 CC AAB51618 represent peptides homologous to Yada, a Yersinia adhesin which
 CC is an important virulence determinant of the Yersinia species. The
 CC invention includes an antibody which binds to the proteobacterial
 CC extracellular peptides, and an immunogenic composition containing the
 CC antibody used as a vaccine to prevent infection by a proteobacteria. The
 CC polypeptides and antibodies are useful in the treatment and prevention of
 CC proteobacterial infections. The polypeptides can also be used to identify
 CC compounds which antagonize the binding of a bacterial adhesion to its
 CC ligand. The host cell can be used to produce the polypeptides in a
 CC suitable culture system. The composition can be used to vaccinate a
 CC patient against a proteobacterial infection
 XX Sequence 24 AA;
 SQ

Query Match 31.7%; Score 33; DB 3; Length 24;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 TAANSKD 20
 |||||
 Db 11 TAANSKD 17

RESULT 10
 AAY49915
 ID AAY49915 standard; peptide; 8 AA.
 XX AAY49915;
 AC AAY49915;
 XX 17-OCT-2003 (revised)
 DT 20-MAR-2003 (revised)
 DT 27-JAN-2000 (first entry)
 XX Gluconobacter suboxydans L-sorbose reductase peptide SEQ ID NO:4.
 DE Gluconobacter suboxydans; L-sorbose reductase; genetic engineering;
 KW mutant.
 KW Gluconobacter oxydans.
 OS AU9920390-A.
 XX 23-SEP-1999.
 PD 11-MAR-1999; 99AU-00020390.
 PF 13-MAR-1998; 98EP-00104546.
 PR (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Hoshino T, Tazoe M, Shinjoh M, Kon T;
 PI WPI; 1999-579276/49.
 DR Set of genetically modified mutants not containing L sorbose reductase.
 XX Example 3; Fig 2; 33pp; English.
 PS The present invention describes a genetically engineered microorganism
 CC derived from a microorganism belonging to the genus Gluconobacter or
 CC Acetobacter which is characterised in that the biological activity for
 CC reducing L-sorbose is substantially nullified by gene recombination. The
 CC present sequence represents a Gluconobacter suboxydans L-sorbose
 CC reductase peptide, used in an example from the present invention for
 CC generating degenerate PCR primers. N.B. This patent is equivalent to the
 CC basic NO9301197 in week 199949. (Updated on 20-MAR-2003 to correct DR
 CC field.) (Updated on 17-OCT-2003 to standardise OS field)
 XX Sequence 8 AA;
 SQ

Query Match 30.8%; Score 32; DB 2; Length 8;
 Best Local Similarity 71.4%; Pred. No. 1.4e+06;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 FPHGMLD 10
 ||:|:|
 Db 1 FPHGMLD 7

RESULT 11

AAB10875
 ID AAB10875 standard; protein; 9 AA.

XX AC AAB10875;
 XX DT 26-JAN-2001 (first entry)
 XX Human 9D7 protein immunogenic fragment SEQ ID NO: 34.
 DE DE
 XX Tumour-associated antigen; 9D7; human; cytostatic; immunogenic;
 KW immunotherapy; cancer; cellular anticancer vaccine; diagnostic; kidney;
 KW lung; colon; breast; carcinoma; Hodgkin's lymphoma.
 XX OS Homo sapiens.
 XX PN DE19909503-Al.
 XX PD 07-SEP-2000.
 XX PF 04-MAR-1999; 99DE-01009503.
 XX PR 04-MAR-1999; 99DE-01009503.

XX PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 XX PI Klade C, Adolf G, Sommergruber W, Heider K;
 XX PS WPI; 2000-588357/56.

XX CC This invention describes a novel tumor-associated antigen, designated 9D7
 CC which has cytostatic activity. The invention also describes a method for
 CC isolating (a) a polypeptide (I) that includes (SI) as part of its
 CC sequence; (b) an immunogenic protein fragment or peptide (II) derived
 CC from 9D7; (c) a pharmaceutical composition containing at least 1 of 9D7,
 CC (I) or (II) as active ingredients; (d) an isolated DNA (III) encoding a
 CC protein (IV) with the immunogenic properties of 9D7 or its fragments; (e)
 CC a recombinant DNA (IIIA) that includes (III); and (f) antibodies (Ab)
 CC directed against 9D7, (I) or (II). 9D7, or its derived immunogenic
 CC peptides, are used to induce a humoral and/or cellular response for use
 CC in vivo or ex vivo immunotherapy of cancer. DNA (III) that encodes 9D7
 CC can be used similarly and cells that express 9D7 are useful in cellular
 CC anticancer vaccines. Antibodies (Ab) specific for 9D7 are useful for
 CC imaging, diagnosing and monitoring cancers, also, when conjugated to
 CC cytotoxins or radionuclide, as therapeutic agents. Peptides derived from
 CC 9D7 may also be used diagnostically to test for an immune response. 9D7-
 CC associated cancers are particularly kidney, lung, colon and breast
 CC carcinoma and Hodgkin's lymphoma. AAB10874-B10899 represent immunogenic
 CC fragments of the human tumor-associated antigen 9D7 which is described in
 CC the method of the invention

XX SQ Sequence 9 AA;

Query Match 30.8%; Score 32; DB 3; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LDLEEIA 15.
 ||:|:|
 Db 1 LDLEEIA 7

Db 2 LDLEEIA 8
 ||:|:|:|

RESULT 12

AAB10876
 ID AAB10876 standard; protein; 10 AA.

XX AC AAB10876;
 XX DT 26-JAN-2001 (first entry)
 XX Human 9D7 protein immunogenic fragment SEQ ID NO: 35.
 DE DE
 XX Tumour-associated antigen; 9D7; human; cytostatic; immunogenic;
 KW immunotherapy; cancer; cellular anticancer vaccine; diagnostic; kidney;
 KW lung; colon; breast; carcinoma; Hodgkin's lymphoma.
 XX OS Homo sapiens.
 XX PN DE19909503-Al.
 XX PD 07-SEP-2000.
 XX PF 04-MAR-1999; 99DE-01009503.
 XX PR 04-MAR-1999; 99DE-01009503.

XX PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 XX PI Klade C, Adolf G, Sommergruber W, Heider K;
 XX PS WPI; 2000-588357/56.

XX CC New tumor-associated antigen 9D7, useful e.g. for immunotherapy of cancer
 CC and for producing diagnostic or therapeutic antibodies.

XX CC Claim 6; Page 37; 50pp; German.

XX CC This invention describes a novel tumor-associated antigen, designated 9D7
 CC which has cytostatic activity. The invention also describes a method for
 CC isolating (a) a polypeptide (I) that includes (SI) as part of its
 CC sequence; (b) an immunogenic protein fragment or peptide (II) derived
 CC from 9D7; (c) a pharmaceutical composition containing at least 1 of 9D7,
 CC (I) or (II) as active ingredients; (d) an isolated DNA (III) encoding a
 CC protein (IV) with the immunogenic properties of 9D7 or its fragments; (e)
 CC a recombinant DNA (IIIA) that includes (III); and (f) antibodies (Ab)
 CC directed against 9D7, (I) or (II). 9D7, or its derived immunogenic
 CC peptides, are used to induce a humoral and/or cellular response for use
 CC in vivo or ex vivo immunotherapy of cancer. DNA (III) that encodes 9D7
 CC can be used similarly and cells that express 9D7 are useful in cellular
 CC anticancer vaccines. Antibodies (Ab) specific for 9D7 are useful for
 CC imaging, diagnosing and monitoring cancers, also, when conjugated to
 CC cytotoxins or radionuclide, as therapeutic agents. Peptides derived from
 CC 9D7 may also be used diagnostically to test for an immune response. 9D7-
 CC associated cancers are particularly kidney, lung, colon and breast
 CC carcinoma and Hodgkin's lymphoma. AAB10874-B10899 represent immunogenic
 CC fragments of the human tumor-associated antigen 9D7 which is described in
 CC the method of the invention

XX SQ Sequence 10 AA;

Query Match 30.8%; Score 32; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LDLEEIA 15
 ||:|:|:|
 Db 1 LDLEEIA 7

RESULT 13

ABP46141

ABP46141 standard; peptide; 14 AA.
 AC ABP46141;
 XX
 DT 19-AUG-2002 (first entry)
 DE Human Blys binding scFv VH CDR3 SEQ ID 2152.
 XX
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.
 XX
 OS Homo sapiens.
 XX
 PN WO200202641-A1.
 XX
 PD 10-JAN-2002.
 XX
 PF 15-JUN-2001; 2001WO-US019110.
 XX
 PR 16-JUN-2000; 2000US-0212210P.
 PR 17-OCT-2000; 2000US-0240816P.
 PR 16-MAR-2001; 2001US-0276248P.
 PR 21-MAR-2001; 2001US-0277379P.
 PR 25-MAY-2001; 2001US-0293499P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
 XX
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
 XX
 DR WPI; 2002-114799/15.
 XX
 PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.
 XX
 PS Claim 2; Page 2940; 3148pp; English.
 XX
 CC This invention describes novel antibodies that immunospecifically bind to
 CC B Lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP4390-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention
 XX
 SQ Sequence 14 AA;
 Query Match 30.8%; Score 32; DB 5; Length 14;
 Best Local Similarity 75.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 FPHGMLDL 11
 Db 7 FPHSLDL 14
 ||| |||
 RESULT 14
 AAO21225
 ID AAO21225 standard; peptide; 19 AA.

XX AAO21225;
 AC
 XX 05-AUG-2002 (first entry)
 DT
 DE Alpha-isomaltosyl-transferring enzymatic activity related peptide #1.
 XX
 KW Alpha-isomaltosyl transferase activity; tetrasaccharide; oligosaccharide;
 KW alpha-isomaltosyl transfer; sugar; glucose polymerisation; viscosity;
 KW calorific value; heat conductivity; sweetness adjuster; taste improver;
 KW texture improver; stabiliser; food additive; drink; cosmetics; drug.
 OS Unidentified.
 XX
 PN WO200240659-A1.
 XX
 PD 23-MAY-2002.
 XX
 PF 16-NOV-2001; 2001WO-JP010044.
 XX
 PR 16-NOV-2000; 2000JP-00350142.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 PI Kubota M, Maruta K, Yamamoto T, Fukuda S;
 XX
 DR WPI; 2002-427088/45.
 XX
 PT Polypeptides of Bacillus origin having alpha-isomaltosyl transferase
 PT activity for manufacture of oligosaccharides as food and drug additives.
 XX
 PS Disclosure; Page 89; 108pp; Japanese.
 XX
 CC The invention relates to polypeptides having alpha-isomaltosyl
 CC transferase activity which are able to synthesise cyclo(1-6)-alpha-D-
 CC glucopyranosyl-(1-3)-alpha-D-glucopyranosyl-(1-6)-alpha-D-glucopyranosyl-
 CC (1-3)-alpha-D-glucopyranosyl) tetrasaccharide by alpha-isomaltosyl
 CC transfer from a sugar having a degree of glucose polymerisation of 3 or
 CC more and having a alpha-(1-6)-glucosyl bond at the non-reducing end and
 CC alpha-(1-4)glucosyl elsewhere. The polypeptides of the invention can be
 CC used in production of oligosaccharides with suitable properties (e.g.
 CC calorific value, heat conductivity, viscosity) for use as sweeteners
 CC adjusters, taste improvers, texture improvers, stabilisers and other
 CC additives for foods, drinks, cosmetics and drugs. This sequence
 CC represents a peptide relating to the alpha-isomaltosyl-transferring
 CC enzymatic activity protein of the invention
 XX
 SQ Sequence 19 AA;
 Query Match 30.8%; Score 32; DB 5; Length 19;
 Best Local Similarity 58.3%; Pred. No. 2.9e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 5 PHGMLDLLEEIAA 16
 Db 8 PYGIDDLVEIOA 19
 ||: ||| |||
 RESULT 15
 ADD73708
 ID ADD73708 standard; peptide; 23 AA.
 XX
 AC ADD73708;
 XX
 DT 29-JAN-2004 (first entry)
 DE
 DE Yeast GlcN-6-P acetyltransferase domain I.
 KW phosphoglucose isomerase activity; ketol-isomerase activity;
 KW acetyltransferase activity; fungal glucose utilisation inhibitor;
 KW antifungal agent; GlcN-6-P acetyltransferase; yeast.
 XX
 OS Synthetic.

OS Saccharomyces cerevisiae.
 XX
 PN US2003194758-A1.
 XX
 PD 16-OCT-2003.
 XX
 PF 10-AUG-2001; 2001US-00927734.
 XX
 PR 10-AUG-2001; 2001US-00927734.
 XX
 PA (MYCO-) MYCOLOGICS INC.
 XX
 PI Selitrennikoff CP, Nakata M;
 XX
 DR WPI; 2003-852592/79.
 XX
 PT Detecting phosphoglucose isomerase, ketol-isomerase and acetyltransferase
 PT activities in a sample, useful for identifying antifungal agents,
 PT comprises measuring the production of Coenzyme A in the sample.
 XX
 PS Disclosure; SEQ ID NO 5; 23pp; English.
 XX
 CC The invention relates to a method of detecting phosphoglucose isomerase,
 CC ketol-isomerase and acetyltransferase activities in a sample which
 CC comprises measuring the production of Coenzyme A. The composition and
 CC methods are useful in identifying fungal glucose utilisation inhibitors
 CC and antifungal agents. The present sequence represents the amino acid
 CC sequence of the yeast GlcN-6-P acetyltransferase domain I.
 XX
 SQ Sequence 23 AA;
 Query Match 30.8%; Score 32; DB 7; Length 23;
 Best Local Similarity 66.7%; Pred.No. 3.6e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 11 LEEIAANSK 19
 :|||
 Db 1 IEDIAVNSK 9

Search completed: July 15, 2004, 14:46:46
 Job time : 43.1333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:40:36 ; Search time 8 53333 Seconds
(without alignments)
225.449 Million cell updates/sec

Title: US-09-171-432A-44
Perfect score: 104
Sequence: 1 KVNPHGMLDLEETAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 6282

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	30.8	26	2	PS0111
2	31	29.8	20	2	B38382
3	31	29.8	27	2	A27108
4	27	26.0	24	2	I53473
5	26	25.0	14	2	S50900
6	26	25.0	15	2	PN0052
7	26	25.0	23	2	A44626
8	26	25.0	25	2	A60619
9	25	24.0	16	2	A54877
10	25	24.0	21	2	A41154
11	25	24.0	26	4	T07533
12	25	24.0	28	2	T06340
13	24	23.1	15	2	P00073
14	24	23.1	17	2	B31435
15	24	23.1	20	2	C60822
16	24	23.1	25	2	PN0632
17	24	23.1	25	2	PC1314
18	24	23.1	26	2	C61590
19	24	23.1	27	2	C44636
20	24	23.1	27	2	JH0755
21	24	23.1	29	2	A35121
22	24	23.1	30	2	S74107
23	24	23.1	30	2	J50645
24	23.5	22.6	29	2	D55998
25	23	22.1	15	2	A54397
26	23	22.1	17	2	S21289
27	23	22.1	19	2	A34233
28	23	22.1	22	2	B26212
29	23	22.1	24	2	PQ0677

30	23	22.1	25	2	T07504
31	23	22.1	25	2	G41606
32	23	22.1	26	2	S07303
33	23	22.1	28	2	I45911
34	23	22.1	29	2	A48427
35	23	22.1	29	2	F85570
36	23	22.1	30	2	F95118
37	22	21.2	10	2	B37196
38	22	21.2	14	1	QMVHMM
39	22	21.2	18	2	A24345
40	22	21.2	20	2	S77981
41	22	21.2	20	2	S46479
42	22	21.2	20	2	A48406
43	22	21.2	20	2	A85659
44	22	21.2	22	2	B32946
45	22	21.2	23	2	PH1726

ALIGNMENTS

RESULT 1

PS0111
beta-galactoside binding lectin - common marmoset (fragments)
C:Species: Callithrix jacchus (common marmoset)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C:Accession: PS0111
R:Ohsawa, F.; Hirano, F.; Natori, S.
J. Biochem. 107, 431-434, 1990
A:Title: Purification and properties of a beta-galactoside-binding lectin from neonatal
A:Reference number: PS0111; MUID:90256719; PMID:1692826
A:Accession: PS0111
A:Molecule type: protein
A:Residues: 1-26 <OHS>

Query Match 30.8%; Score 32; DB 2; Length 26;
Best Local Similarity 53.3%; Pred. No. 95;
Matches 8; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 6 HGMLDLEETAANSKD 20
||| | | | | | | |
DB 4 HG--DANTIVXNSKD 16

RESULT 2

B38382
15K protein B - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 02-Sep-2000
C:Accession: B38382
R:Ooi, C.E.; Weiss, J.; Levy, O.; Elsbach, P.
J. Biol. Chem. 265, 15956-15962, 1990
A:Title: Isolation of two isoforms of a novel 15-kDa protein from rabbit polymorphonucle
A:Reference number: A38382; MUID:90368817; PMID:2203792
A:Accession: B38382
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <OOI>
C:Superfamily: cathelin; cystatin homology

Query Match 29.8%; Score 31; DB 2; Length 20;
Best Local Similarity 42.9%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 PHGMLDLEETAANS 18
||| | | | | | | |
DB 2 PHRLRYEEVVAQA 15

RESULT 3

A27108
larval-specific lipoprotein - honeybee (fragment)
C:Species: Apis mellifera (honeybee)

C;Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 31-Dec-1993

C;Accession: A27108

R;Shipman, B.A.; Ryan, R.O.; Schmidt, J.O.; Law, J.H.

Biochemistry 26, 1885-1889, 1987

A;Title: Purification and properties of a very high density lipoprotein from the hemolysate of sheep erythrocytes

A;Reference number: A27108; MUID:87242376; PMID:3109474

A;Accession: A27108

A;Molecule type: protein

A;Residues: 1-27 <SHI>

C;Keywords: lipid binding; lipoprotein

Query Match 29.8%; Score 31; DB 2; Length 27;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 FPHGMLDLEIAANSK 19

||||| : : : : :
Db 4 FPHGKLVTKYIADVK 19

RESULT 4

153473 phosphoprotein phosphatase (EC 3.1.3.16) 2A catalytic chain, clone H6 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 22-Jun-1999

C;Accession: 153473

R;Guen, M.X.; Chen, Y.H.; Cohen, P.T.

FEBS Lett. 306, 54-58, 1992

A;Title: Polymerase chain reactions using Saccharomyces, Drosophila and human DNA predicted

A;Reference number: 153473; MUID:92331787; PMID:1321058

A;Accession: 153473

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-24 <RES>

A;Cross-references: GB:S40009; NID:G251185; PIDN:AA22470.1; PID:G251186

C;Superfamily: phosphoprotein phosphatase; phosphoesterase core homology; phosphoprotein

C;Keywords: iron; phosphoric monoester hydrolase; serine/threonine-specific phosphatase;

Query Match 26.0%; Score 27; DB 2; Length 24;

Best Local Similarity 33.3%; Pred. No. 6e+02;

Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 6 HGMLEIEIAANSKD 20

||||| : : : : :
Db 3 HDLWELFRIGGKSPD 17

RESULT 5

S50900

chlorophyll a/b-binding protein Lhcb5 - spinach (fragment)

N;Alternate names: light-harvesting complex LHCIIc protein

C;Species: Spinacia oleracea (spinach)

C;Date: 19-Mar-1997 #sequence_revision 23-Apr-1999 #text_change 23-Apr-1999

C;Accession: S50900

R;Walters, R.G.; Ruban, A.V.; Horton, P.

Eur. J. Biochem. 226, 1063-1069, 1994

A;Title: Higher plant light-harvesting complexes LHCIIa and LHCIIc are bound by dicyclohexyl

A;Reference number: S50900; MUID:95112835; PMID:7813461

A;Accession: S50900

A;Molecule type: protein

A;Residues: 1-14 <WAL>

C;Superfamily: chlorophyll a/b-binding protein

C;Keywords: chlorophyll; chloroplast; light-harvesting complex; photosynthesis; photosyn

Query Match 25.0%; Score 26; DB 2; Length 14;

Best Local Similarity 44.4%; Pred. No. 4.7e+02;

Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 PHGMLDLEE 13

||||| : : : : :
Db 4 PDGLDRDD 12

RESULT 6

PN0052

pyruvate kinase (EC 2.7.1.40) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 29-Oct-1997 #sequence_revision 23-Jan-1998 #text_change 23-Jan-1998

C;Accession: PN0052

R;Kato, H.

Kawasaki Igakkaishi 22, 245-259, 1996

A;Title: Analysis of proteins isolated by two dimensional electrophoresis of mouse neuro

A;Reference number: PN0041

A;Accession: PN0052

A;Molecule type: protein

A;Residues: 1-15 <KAT>

A;Experimental source: neuroblastoma cell

C;Comment: The molecular mass is 40,500 and the pI is 5.48.

C;Keywords: brain; phosphotransferase

Query Match 25.0%; Score 26; DB 2; Length 15;

Best Local Similarity 38.5%; Pred. No. 5.1e+02;

Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLEE 14

||||| : : : : :
Db 3 VNLPGAADLPAV 15

RESULT 7

A44626

dihydropyrimidine dehydrogenase (NADP) (EC 1.3.1.2) - bovine (fragment)

N;Alternate names: dihydrothymine dehydrogenase; dihydouracil dehydrogenase (NADP+)

C;Species: Bos primigenius taurus (cattle)

C;Date: 02-Aug-1994 #sequence_revision 23-Mar-1995 #text_change 03-Jun-2002

C;Accession: A44626; A41066

R;Porter, D.J.T.; Chestnut, W.G.; Merrill, B.M.; Spector, T.

J. Biol. Chem. 267, 5236-5242, 1992

A;Title: Mechanism-based inactivation of dihydropyrimidine dehydrogenase by 5-ethynylur

A;Reference number: A44626; MUID:92184771; PMID:1544906

A;Accession: A44626

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-23 <POR>

R;Porter, D.J.T.; Chestnut, W.G.; Taylor, L.C.E.; Merrill, B.M.; Spector, T.

J. Biol. Chem. 266, 19988-19994, 1991

A;Title: Inactivation of dihydropyrimidine dehydrogenase by 5-iodouracil.

A;Reference number: A41066; MUID:32041818; PMID:1939061

A;Accession: A41066

A;Molecule type: protein

A;Residues: 13-23 <PO2>

C;Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h

C;Keywords: NADP; oxidoreductase

F;16/Binding site: substrate (Cys) #status experimental

Query Match 25.0%; Score 26; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 8.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PHGM 8

||||| : : : : :
Db 17 PHGM 20

RESULT 8

A60619

prolactin - Atlantic salmon (fragment)

C;Species: Salmo salar (Atlantic salmon)

C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993

C;Accession: A60619

R;Andersen, O.; Skibeli, V.; Gautvik, K.M.

Gen. Comp. Endocrinol. 73, 354-360, 1989

A;Title: Purification and characterization of Atlantic salmon prolactin.

A;Reference number: A60619; MUID:89171973; PMID:2925076

A;Accession: A60619

A;Molecule type: protein

A;Residues: 1-25 <AND>
C;Superfamily: prolactin
C;Keywords: hormone; pituitary

Query Match 25.0%; Score 26; DB 2; Length 25;
Best Local Similarity 42.9%; Pred. No. 9.2e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 GMLDLEETAAANSKD 20
| : ||| | : |
Db 2 GLSDLMERASQRS 15

RESULT 9

A54877
alpha-conotoxin PnIA [validated] - cone shell (Conus pennaceus)

N;Alternate names: alpha-Ctx-PnIA
C;Species: Conus pennaceus
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000
C;Accession: A54877
R;Painzliber, M.; Hasson, A.; Oren, R.; Burlingame, A.L.; Gordon, D.; Spira, M.E.; Zlotk
Biochemistry 33, 9523-9525, 1994
A;Title: New mollusc-specific alpha-conotoxins block Aplysia neuronal acetylcholine rece
A;Reference number: A54877; MUID:94347719; PMID:8068627
A;Accession: A54877
A;Molecule type: protein
A;Residues: 1-16 <FAI>
R;Hu, S.H.; Gehrmann, J.; Guddat, L.W.; Alewood, P.F.; Craik, D.J.; Martin, J.L.
submitted to the Brookhaven Protein Data Bank, January 1996
A;Reference number: A66355; PDB:1PEN
A;Contents: annotation; X-ray crystallography, 1.1 angstroms; residues 1-16
C;Comment: This alpha-conotoxin, as an acetylcholine receptor inhibitor, is a postsynapt
C;Superfamily: alpha-conotoxin
C;Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; postsynaptic neuro
F;2-8,3-16/disulfide bonds: #status experimental
F;16/Modified site: amidated carboxyl end (Cys) #status experimental

Query Match 24.0%; Score 25; DB 2; Length 16;
Best Local Similarity 42.9%; Pred. No. 8.1e+02;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 7 GMLDLEETAAANSKD 20
| : ||| | : |
Db 1 GCCSLPPCAANNPD 14

RESULT 10

A41154
ethylene-forming enzyme - tomato (fragment)
C;Species: Lycopersicon esculentum (tomato)
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 11-Jun-1999
C;Accession: A41154
R;Hamilton, A.J.; Bouzayen, M.; Grierson, D.
Proc. Natl. Acad. Sci. U.S.A. 88, 7434-7437, 1991
A;Title: Identification of a tomato gene for the ethylene-forming enzyme by expression i
A;Reference number: A41154; MUID:91334480; PMID:1714605
A;Accession: A41154
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-21 <HAM>
A;Cross-references: GB:S50262; NID:g233722; PIDN:AAB19492.1; PID:g233723
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 24.0%; Score 25; DB 2; Length 21;
Best Local Similarity 29.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 7; Mismatches 3; Indels 2; Gaps 1;

QY 3 NPPHGMGLDLEETAAANSK 19
||| : : : ||| : :
Db 3 NFP--IINLEKLNQNGER 17

RESULT 11

T07533

Probable photosystem II D2 protein pseudogene - Japanese black pine chloroplast
C;Species: chloroplast Pinus thunbergiana (Japanese black pine)
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 08-Dec-2000
C;Accession: T07533

R;Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A;Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genome

A;Reference number: Z16030; MUID:95024047; PMID:7937893

A;Accession: T07533

A;Status: translated from GB/EMBL/DBJ; conceptual translation of pseudogene

A;Molecule type: DNA

A;Residues: 1-26 <WAK>

A;Cross-references: EMBL:D17510; NID:g529643; PIDN:BAA04409.1; PID:g1262694

C;Genetics:

A;Genome: chloroplast

C;Keywords: chloroplast; pseudogene

Query Match 24.0%; Score 25; DB 4; Length 26;
Best Local Similarity 38.5%; Pred. No. 1.4e+03;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 NPPHGMGLDLEETA 15
||| : : ||| :
Db 11 NFPLDLAAVESIS 23

RESULT 12

T06340
ribosomal protein S16 - soybean chloroplast (fragment)
C;Species: chloroplast Glycine max (soybean)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
C;Accession: T06340
R;Nielsen, N.C.
submitted to the EMBL Data Library, May 1995

A;Reference number: Z15613

A;Accession: T06340

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-28 <NIE>

A;Cross-references: EMBL:U26948; NID:g984307; PIDN:AAA80642.1; PID:g984308

A;Experimental source: cultivar Resnik; leaf

C;Genetics:

A;Gene: rps16

A;Genome: chloroplast

C;Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 24.0%; Score 25; DB 2; Length 28;
Best Local Similarity 83.3%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 MLDLEE 13
| : ||| | : |
Db 10 MFDLEE 15

RESULT 13

PQ0073

T-cell receptor beta chain (BTB48) - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995

C;Accession: PQ0073

R;Tanaka, A.; Ishiguro, N.; Shinagawa, M.

submitted to JIPID, May 1990

A;Description: Sequence analysis of bovine T-cell receptor beta chain genes.

A;Reference number: JQ0472

A;Accession: PQ0073

A;Molecule type: mRNA

A;Residues: 1-15 <TAN>

A;Experimental source: T cell

C;Genetics:

A;Gene: BTB48

C;Keywords: receptor

Query Match 23.1%; Score 24; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NFPHG 7
|||
Db 5 NFGHG 9

RESULT 14

B31435
adherence lectin light chain - Entamoeba histolytica (fragment)
C; Species: Entamoeba histolytica
C; Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 30-Sep-1993
C; Accession: B31435
R; Petri Jr., W.A.; Chapman, M.D.; Snodgrass, T.; Mann, B.J.; Broman, J.; Ravdin, J.I.
J. Biol. Chem. 264, 3007-3012, 1989
A; Title: Subunit structure of the galactose and N-acetyl-D-galactosamine-inhibitable ad
A; Reference number: A31435; MUID:89123408; PMID:2536731
A; Accession: B31435
A; Status: preliminary
A; Molecule type: protein
A; Residues: 1-17 <PET>

Query Match 23.1%; Score 24; DB 2; Length 17;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 NFPHG 7
|||
Db 12 NPYG 16

RESULT 15

C60822
cytochrome P450 UT501 - rat (fragment)
C; Species: Rattus norvegicus (Norway rat)
C; Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C; Accession: C60822
R; Ameliazad, Z.; Marbonne, J.F.; Wolf, C.R.; Robertson, L.W.; Oesch, F.
Biochem. Pharmacol. 37, 3245-3249, 1988
A; Title: Effect of nutritional imbalances on cytochrome P-450 isozymes in rat liver.
A; Reference number: A60822; MUID:88293549; PMID:3041969
A; Accession: C60822
A; Status: preliminary
A; Molecule type: protein
A; Residues: 1-20 <AME>

Query Match 23.1%; Score 24; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 9 LDLEEIAA 16
:|:|:|
Db 1 MDLADVAA 8

Search completed: July 15, 2004, 14:52:13
Job time : 10.5333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:01 ; Search time 5.3333 Seconds
(without alignments)
195.263 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104

Sequence: 1 KVNPHGMLDLEETAAANSK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 2073

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	29.8	27	1	LSP_APIME
2	26	25.0	20	1	FIBB_SHEEP
3	25	24.0	16	1	CXAA_CONPE
4	24	23.1	12	1	PR2B_VIPBO
5	24	23.1	22	1	CRT_CLOPA
6	24	23.1	23	1	CLVB_STYCL
7	24	23.1	30	1	DIU2_HYLLI
8	24	23.1	30	1	DIU2_MANSE
9	23.5	22.6	29	1	BR2D_RANES
10	23	22.1	19	1	ITHA_PERAM
11	23	22.1	20	1	PGK_CLOPA
12	23	22.1	22	1	23KD_BACST
13	23	22.1	26	1	RBL_VICFA
14	23	22.1	29	1	SODC_OLEEU
15	22	21.2	10	1	BPP2_BOTIN
16	22	21.2	14	1	MAST_VESNA
17	22	21.2	20	1	COXA_THUOB
18	22	21.2	21	1	GRX_FUNCN
19	22	21.2	29	1	GALA_ONCMY
20	22	21.2	29	1	RS7_METTE
21	21.5	20.7	21	1	SVR_RAT
22	21	20.2	15	1	GLN2_PINDS
23	21	20.2	17	1	RS13_PARDE
24	21	20.2	19	1	DHAB_COMTE
25	21	20.2	21	1	TRYP_APIMS
26	21	20.2	28	1	ARYC_NOCGL
27	21	20.2	28	1	VI03_VACCP
28	20	19.2	8	1	AKHG_GRYBI
29	20	19.2	9	1	RE42_LITRU
30	20	19.2	10	1	HTF_NAUCI
31	20	19.2	13	1	ACT7_SOYBN
32	20	19.2	13	1	CD71_LITETW
33	20	19.2	14	1	MAST_VESLE

34	20	19.2	15	1	CIQA_RAT	P31720 rattus norv
35	20	19.2	15	1	SODP_PINDS	P81082 pinus pinas
36	20	19.2	16	1	AU21_LITRA	P82388 litoria ran
37	20	19.2	17	1	ACT6_SOYBN	P15986 glycine max
38	20	19.2	19	1	UP25_UPEIN	P82031 uperoleia i
39	20	19.2	21	1	PEDB_HYDAT	P80577 hydra atten
40	20	19.2	23	1	VG22_BPT2	P21596 bacterioph
41	20	19.2	23	1	VG22_BPT6	P21597 bacterioph
42	20	19.2	24	1	HS9B_RABIT	P30947 oryctolagus
43	20	19.2	24	1	RSS_VIBPR	P52856 vibrio prot
44	20	19.2	25	1	ALR_PSEFL	P33162 pseudomonas
45	20	19.2	25	1	AU51_LITRA	P82401 litoria ran

ALIGNMENTS

RESULT 1						
LSP_APIME						
ID	LSP_APIME	STANDARD;	PRT;	27	AA.	
AC	P09355;					
DT	01-MAR-1989 (Rel. 10, Created)					
DT	01-MAR-1989 (Rel. 10, Last sequence update)					
DT	15-MAR-2004 (Rel. 43, Last annotation update)					
DE	Larval-specific very high density lipoprotein (VHDL) (Fragment).					
OS	Apis mellifera (Honeybee).					
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Prerygota;					
OC	Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;					
OC	Apidae; Apis.					
OX	NCBI_TaxID=7460;					
RN	[1]					
RP	SEQUENCE.					
RC	TISSUE=Hemolymph;					
RX	MEDLINE=87242376; PubMed=3109474;					
RA	Shipman B.A., Ryan R.O., Schmidt J.O., Law J.H.;					
RT	"Purification and properties of a very high density lipoprotein from the hemolymph of the honeybee Apis mellifera.";					
RL	Biochemistry 26:1885-1889(1987).					
CC	!- FUNCTION: Unknown (it might play a role in lipid transport and/or storage protein metabolism during metamorphosis).					
CC	!- SUBUNIT: Homodimer.					
CC	!- SUBCELLULAR LOCATION: Extracellular.					
CC	!- TISSUE SPECIFICITY: Hemolymph.					
CC	!- DEVELOPMENTAL STAGE: Vhdl is present in high amounts in hemolymph only at the end of larval life.					
DR	PIR; A27108; A27108.					
KW	Lipoprotein; Hemolymph.					
FT	NON TER 27					
SQ	SEQUENCE 27 AA; 2885 MW; A9DDF7A75A65D10D CRC64;					
Query Match 29.8%; Score 31; DB 1; Length 27;						
Best Local Similarity 43.8%; Pred. NO. 88;						
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;						
QY	4	FPHGMLDLEETAAANSK 19				
Db	4	FPHGKLVITYKIADVK 19				
RESULT 2						
FIBB_SHEEP						
ID	FIBB_SHEEP	STANDARD;	PRT;	20	AA.	
AC	F14470;					
DT	01-JAN-1990 (Rel. 13, Created)					
DT	01-JAN-1990 (Rel. 13, Last sequence update)					
DT	10-OCT-2003 (Rel. 42, Last annotation update)					
DE	Fibrinogen beta chain [Contains: Fibrinopeptide B] (fragment).					
GN	FGB.					
OS	Ovis aries (Sheep), and					
OS	Capra hircus (Goat).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;					
OC	Bovidae; Caprinae; Ovis.					


```
QY 7 GMLDLEE 13
   ||: |||
Db 6 GMMILEE 12

RESULT 5
ID CRT_CLOPA STANDARD; PRT; 22 AA.
AC P81357;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 3-hydroxybutyryl-CoA dehydratase (EC 4.2.1.55) (Crotonase) (CP 24)
DE (Fragment).
GN CRT.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]_SEQUENCE.
RP STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flensburg R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- CATALYTIC ACTIVITY: (3R)-3-hydroxybutanoyl-CoA = crotonoyl-CoA +
H(2)O.
CC -!- PATHWAY: Butyrate/butanol-producing pathway.
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
PROTEIN IS: 5.2, ITS MW IS: 30.9 kDa.
CC -!- SIMILARITY: Belongs to the enoyl-CoA hydratase/isomerase family.
DR InterPro; IPR001753; ENCOA_hydratase.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; PARTIAL.
KW Fatty acid metabolism; Enzyme N -> D.
FT VARIANT 20 20
FT NON TER 22 22
SQ SEQUENCE 22 AA; 2296 MW; 95CD79FAD3FE7CDB CRC64;

Query Match 23.1%; Score 24; DB 1; Length 22;
Best Local Similarity 40.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLEIAA 16
   ||: |||
Db 1 MGFNGGILEGEEXA 15

RESULT 6
ID CLVB_STYCL STANDARD; PRT; 23 AA.
AC F80711;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Clavadin B.
OS Styela clava (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Styela.
OX NCBI_TaxID=7725;
RN [1]_SEQUENCE.
RP TISSUE=Hemocyte;
RX MEDLINE=97154671; PubMed=9001389;
RA Lee I.H., Zhao C., Cho Y., Harwig S.S.L., Cooper E.L., Lehrer R.I.;
RT "Clavamins, alpha-helical antimicrobial peptides from tunicate
hemocytes.";
RL FEBS Lett. 400:158-162(1997).
CC -!- FUNCTION: Has antimicrobial activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
DR InterPro; IPR008453; Clavamin.
DR Pfam; PF05452; Clavamin; 1.

KW Antibiotic; Amidation.
FT MOD_RES 23 23
SQ SEQUENCE 23 AA; 2695 MW; B30E236441195B15 CRC64;

Query Match 23.1%; Score 24; DB 1; Length 23;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NPPHG 7
   |||
Db 14 NFVHG 18

RESULT 7
ID DIU2_HYLLI STANDARD; PRT; 30 AA.
AC P82015;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Diuretic hormone 2 (DH-2) (Diuretic peptide 2) (DP-2) (DH(30)).
OS Hyles lineata (Whiteline sphinx moth).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;
OC Sphingidae; Macroglossinae; Hyles.
OX NCBI_TaxID=103890;
RN [1]_SEQUENCE.
RP SEQUENCE, ACTIVITY, AND AMIDATION.
RC TISSUE=Head;
RX MEDLINE=20161578; PubMed=10696588;
RA Furuya K., Harper M.A., Schegg K.M., Shooley D.A.;
RT "Isolation and characterization of CRF-related diuretic hormones from
the whiteline sphinx moth Hyles lineata.";
RL Insect Biochem. Mol. Biol. 20:127-133(2000).
CC -!- FUNCTION: Regulation of fluid secretion. May stimulate primary
urine secretion by Malpighian tubules and causes a dose-dependent
stimulation of cAMP levels in the tubules.
CC -!- SURCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=3574; METHOD=MALDI.
CC -!- SIMILARITY: Belongs to the sauvagine/corticotropin-releasing
factor/urotensin I family.
DR InterPro; IPR000187; corticoliberin.
DR Pfam; PF00473; CRF; 1.
DR SMART; SM00039; CRF; 1.
DR PROSITE; PS00511; CRF; FALSE_NEG.
KW Hormone; Amidation.
FT MOD_RES 30 30
SQ SEQUENCE 30 AA; 3575 MW; 5C2D6BD2D8BFC67 CRC64;

Query Match 23.1%; Score 24; DB 1; Length 30;
Best Local Similarity 30.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 11 LEEIAANSKD 20
   ||: |||
Db 16 MEKVAQNRRN 25

RESULT 8
ID DIU2_MANSE STANDARD; PRT; 30 AA.
AC P24858;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Diuretic hormone 2 (DH-2) (Diuretic peptide 2) (DP-2) (DPIL).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;
OC Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]_SEQUENCE.
```

```

RC TISSUE=Corpora cardiaca;
RX MEDLINE=92109782; PubMed=1764106;
RA Blackburn M.B., Kingan T.G., Bodnar W., Shabanowitz J., Hunt D.F.,
RA Kempe T., Wagner R.M., Raina A.K., Schnee M.E., Ma M.C.;
RT "Isolation and identification of a new diuretic peptide from the
RT tobacco hornworm, Manduca sexta.";
RL Biochem. Biophys. Res. Commun. 181:927-932(1991).
CC -!- FUNCTION: Regulation of fluid secretion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the sauvagine/corticotropin-releasing
CC factor/urotensin I family.
DR PIR: JS0645; JS0645.
DR InterPro: IPR000187; corticoliberin.
DR Pfam: PF00473; CRF, 1.
DR SMART: SM00039; CRF; 1.
DR PROSITE: PS00511; CRF; FALSE_NEG.
KW Hormone; Amidation.
FT MOD RES 30 30
FT SEQUENCE 30 AA; 3561 MW; 5C2D6BD2DD8A2717 CRC64;
SQ SEQUENCE 30 AA; 3561 MW; 5C2D6BD2DD8A2717 CRC64;

Query Match 23.1%; Score 24; DB 1; Length 30;
Best Local Similarity 30.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 11 LEEIAANSKD 20
DB 16 MEKVAQNRRN 25

RESULT 9
BR2D_RANES
ID BR2D_RANES STANDARD; PRT; 29 AA.
AC P40840;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Brevinin-2ed.
OS Rana esculenta (Edible frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8401;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=94216303; PubMed=8163497;
RA Sirmaco M., Mignogna G., Barra D., Bossa F.;
RT "Antimicrobial peptides from skin secretions of Rana esculenta.
RT Molecular cloning of cDNAs encoding esculentin and brevinins and
RT isolation of new active peptides.";
RL J. Biol. Chem. 269:11956-11961(1994).
CC -!- FUNCTION: Shows antibacterial activity against representative
CC Gram-negative and Gram-positive bacterial species, and hemolytic
CC activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: Belongs to the frog skin active peptide (FSAP) family.
CC Brevinin subfamily.
DR PIR: D55998; D55998.
KW Amphibian defense peptide; Antibiotic; Hemolysis.
FT DISULFID 23 29
FT SEQUENCE 29 AA; 3002 MW; 33299AFDA6CE8612 CRC64;
SQ SEQUENCE 29 AA; 3002 MW; 33299AFDA6CE8612 CRC64;

Query Match 22.6%; Score 23.5; DB 1; Length 29;
Best Local Similarity 46.2%; Pred. No. 1.8e+03;
Matches 6; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 7 GMLD-LEEIAANS 18
DB 1 GILDSLKLAKNA 13

RESULT 10
ITHA_PERAM

```

```

ID ITHA_PERAM STANDARD; PRT; 19 AA.
AC P19986;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Trehalase inhibitor (Fragment).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattelloidea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph; PubMed=2777784;
RX MEDLINE=89380218;
RA Hayakawa Y., Jahagirdar A.P., Yaguchi M., Downer R.G.H.;
RT "Purification and characterization of trehalase inhibitor from
RT hemolymph of the American cockroach, Periplaneta americana.";
RL J. Biol. Chem. 264:16165-16169(1989).
CC -!- FUNCTION: Under resting conditions, inhibits trehalase in a
CC dose-dependent manner.
CC PIR: A34233; A34233.
KW Glycoprotein.
FT NON TER 19 19
FT SEQUENCE 19 AA; 2016 MW; A60C3A3397AF9A22 CRC64;
SQ SEQUENCE 19 AA; 2016 MW; A60C3A3397AF9A22 CRC64;

Query Match 22.1%; Score 23; DB 1; Length 19;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KVNPPHGMLD 10
DB 9 KXVDPGALN 18

RESULT 11
PGK_CLOPA
ID PGK_CLOPA STANDARD; PRT; 20 AA.
AC P81346;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative phosphoglycerate kinase (EC 2.7.2.3) (CP 11) (Fragment).
GN PGK.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -!- PATHWAY: Second phase of glycolysis; second step.
CC -!- SUBUNIT: Monomer (by similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC 5.6, ITS MW IS: 56.2 kDa.
CC -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
DR HANAP; MF_00145; -.
DR InterPro: IPR001576; PGK.
DR PROSITE: PS00111; PGLYCERATE_KINASE; PARTIAL.
KW Transferase; Kinase; Glycolysis.
FT NON TER 20 20
FT SEQUENCE 20 AA; 2355 MW; 749E31E595C85529 CRC64;
SQ SEQUENCE 20 AA; 2355 MW; 749E31E595C85529 CRC64;

Query Match 22.1%; Score 23; DB 1; Length 20;
Best Local Similarity 28.6%; Pred. No. 1.5e+03;

```


Matches 4; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 6 HGMLDLEIAANSK 19
Db 2 YNLLTIEDIVKKG 15

RESULT 12
23KD BACST STANDARD; PRT; 22 AA.
ID P80166;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE 23 kDa basic protein (Fragment).
OS Bacillus stearotherophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE.
RA Vorgias C.E.;
RL Submitted (Oct-1992) to Swiss-Prot.
DR InterPro; IPR007551; DUF520.
DR Pfam; PF04461; DUF520: 1.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2376 MW; 96C604E42CE0BF6C CRC64;

Query Match 22.1%; Score 23; DB 1; Length 22;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 LDLEIEIA 15
Db 12 VDLSEVA 18

RESULT 13
RBL_VICFA STANDARD; PRT; 26 AA.
ID P05699;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ribulose biphosphate carboxylase large chain precursor (EC 4.1.1.39)
DE (RuBisCO large subunit) (Fragment).
GN RBCL.
OS Vicia faba (Broad bean).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
OX NCBI_TaxID=3906;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinozaki K., Sun C.-R., Sugiura M.;
RT "Gene organization of chloroplast DNA from the broad bean Vicia faba.";
RL Mol. Gen. Genet. 197;363-367(1984).
CC -!- FUNCTION: RuBisCO catalyzes two reactions: the carboxylation of D-ribulose 1,5-bisphosphate, the primary event in photosynthetic carbon dioxide fixation, as well as the oxidative fragmentation of the pentose substrate in the photorespiration process. Both reactions occur simultaneously and in competition at the same active site.
CC -!- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + CO(2) + H(2)O = 2 3-phospho-D-glycerate.
CC -!- CATALYTIC ACTIVITY: D-ribulose 1,5-bisphosphate + O(2) = 3-phospho-D-glycerate + 2-phosphoglycolate.
CC -!- SUBUNIT: 8 large chains + 8 small chains.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: Belongs to the RuBisCO large chain family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01167; CAA25614.1; -.
DR PIR; S07303; S07303.
DR InterPro; IPR000685; RuBisCO_large.
DR PROSITE; PS00157; RUBISCO LARGE; PARTIAL.
KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation.
FT PROPEP 1 2
FT CHAIN 3 >26
FT RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE
FT CHAIN.
FT MOD_RES 3 3
FT NON_TER 26 26
FT ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 26 AA; 2955 MW; 754F486180B67E98 CRC64;

Query Match 22.1%; Score 23; DB 1; Length 26;
Best Local Similarity 50.0%; Pred. No. 1.9e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KVNPHGMLD 10
Db 10 KVGFGAGVD 19

RESULT 14
SODC_OLIEU STANDARD; PRT; 29 AA.
ID P80740;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) (Allergen Ole e 5) (Ole e
DE V) (Fragment).
OS Olea europaea (Common olive).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamids; Lamiales; Oleaceae; Olea.
OX NCBI_TaxID=4146;
RN [1]
RP SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=98160390; PubMed=9500754;
RA Boluda L., Alonso C., Fernandez-Caldas E.;
RT "Purification, characterization, and partial sequencing of two new
RT allergens of Olea europaea".
RL J. Allergy Clin. Immunol. 101:210-216(1998).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
CC similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; sodcu; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; PARTIAL.
DR PROSITE; PS00332; SOD_CU_ZN_2; PARTIAL.
KW Antioxidant; Oxidoreductase; Metal-binding; Copper; Zinc; Allergen.
FT NON_TER 29 29
SQ SEQUENCE 29 AA; 2973 MW; 836C7A193EDAD71E CRC64;

Query Match 22.1%; Score 23; DB 1; Length 29;
Best Local Similarity 60.0%; Pred. No. 2.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 PHGML 9
Db 13 PHGIV 17

```

RESULT 15
BPP2 BOTIN
ID _BPP2 BOTIN STANDARD; PRT; 10 AA.
AC P30422;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bradykinin-potentiating peptide S4,3,1 (10C) (Angiotensin-converting
enzyme inhibitor).
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -!- FUNCTION: This peptide both inhibits the activity of the
CC angiotensin-converting enzyme and enhances the action of
CC bradykinin by inhibiting the kinases that inactivate it.
CC It acts as an indirect hypotensive agent.
CC PIR; B37196; B37196.
DR Hypotensive agent; Pyrrolidone carboxylic acid.
KW MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT SEQUENCE 10 AA; 1213 MW; 30C53546C761F773 CRC64;
SQ
Query Match 21.2%; Score 22; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NPPH 6
|:|
Db 2 NPPH 5

Search completed: July 15, 2004, 14:47:33
Job time : 6.33333 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:31 ; Search time 25.7333 Seconds

(without alignments)
245.221 Million cell updates/sec

Title: US-09-171-432A-44
Perfect score: 104
Sequence: 1 KVFPHGMLDLEETAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 18020

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	28.8	13	11 Q80XX0	Q80xx0 cricetulus
2	30	28.8	19	5 Q9TWG5	Q9twg5 latrodectus
3	30	28.8	22	4 Q96Q47	Q96q47 homo sapien
4	29	27.9	21	2 P70861	P70861 borrelia bu
5	28.5	27.4	16	8 Q8SLG8	Q8slg8 taraxacum (
6	28.5	27.4	16	8 Q8SLF6	Q8slf6 taraxacum (
7	28.5	27.4	16	8 Q9SA19	Q9sa19 jurinea cya
8	28.5	27.4	16	8 Q8SLE2	Q8sle2 taraxacum s
9	28.5	27.4	16	8 Q95AY0	Q95ay0 artemisia r
10	28.5	27.4	16	8 Q8SLF3	Q8slf3 taraxacum (
11	28.5	27.4	17	8 Q8SLE1	Q8sle1 taraxacum (
12	28.5	27.4	17	8 Q8SLE8	Q8sle8 taraxacum (
13	28.5	27.4	17	8 Q8SLE0	Q8sle0 taraxacum (
14	28.5	27.4	17	8 Q8SLG4	Q8slg4 taraxacum (
15	28.5	27.4	17	8 Q8SLG0	Q8slg0 taraxacum (
16	28.5	27.4	17	8 Q8SLG1	Q8slg1 taraxacum (

17	28.5	27.4	18	8 Q8SLF2	Q8slf2 taraxacum (
18	28.5	27.4	18	8 Q8SLG9	Q8slg9 taraxacum s
19	28.5	27.4	18	8 Q8SLE9	Q8sle9 taraxacum (
20	28.5	27.4	18	8 Q8SLE3	Q8sle3 taraxacum (
21	28.5	27.4	19	8 Q8SLF8	Q8slf8 taraxacum (
22	28.5	27.4	19	8 Q8SLG6	Q8slg6 taraxacum (
23	28.5	27.4	19	8 Q8SLE4	Q8sle4 taraxacum (
24	28.5	27.4	19	8 Q8SLF1	Q8slf1 taraxacum s
25	28.5	27.4	19	8 Q8SLD9	Q8slld9 taraxacum s
26	28.5	27.4	19	8 Q8SLE5	Q8sle5 taraxacum (
27	28.5	27.4	19	8 Q8SLF9	Q8slf9 taraxacum (
28	28.5	27.4	19	8 Q8SLG3	Q8slg3 taraxacum (
29	28.5	27.4	19	8 Q8SLG2	Q8slg2 taraxacum (
30	28.5	27.4	19	8 Q8SLE7	Q8sle7 taraxacum (
31	28.5	27.4	19	8 Q8SLE6	Q8sle6 taraxacum (
32	28.5	27.4	22	8 Q9MSX4	Q9msx4 cynara humi
33	28.5	27.4	22	8 Q9MSX2	Q9msx2 gallactites
34	28.5	27.4	22	8 Q9MSW6	Q9msw6 silybum mar
35	28.5	27.4	22	8 Q9MSW9	Q9msw9 onopordum a
36	28.5	27.4	22	8 Q9MSW5	Q9msw5 tyrimnus le
37	28.5	27.4	22	8 Q9MSX6	Q9msx6 cousinia hy
38	28.5	27.4	22	8 Q9MSW7	Q9msw7 ptilostemon
39	28.5	27.4	22	8 Q9MSX7	Q9msx7 cirsium vul
40	28.5	27.4	22	8 Q9MSX0	Q9msx0 notobasis s
41	28.5	27.4	22	8 Q9MSW8	Q9msw8 picnomon ac
42	28.5	27.4	22	8 Q9MSX3	Q9msx3 echinops sp.
43	28.5	27.4	22	8 Q9MSX9	Q9msx9 arctium lap
44	28.5	27.4	23	8 Q8SLH1	Q8slh1 taraxacum (
45	28.5	27.4	23	8 Q8SLH0	Q8slh0 taraxacum (

ALIGNMENTS

RESULT 1
ID Q80XX0 PRELIMINARY; PRT; 13 AA.
AC Q80XX0
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypoxanthine phosphoribosyl transferase (Fragment).
GN HPRT.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]_TaxID=10029;
RP SEQUENCE FROM N.A.
RX MEDLINE=93024555; PubMed=1383700;
RA Fuscoe J.C.; Zimmerman L.J.; Fekete A.; Setzer R.W.; Rossiter B.J.;
RT "Analysis of X-ray-induced HPRT mutations in CHO cells: insertion and
RT deletions.";
RL Mutat. Res. 269:171-183(1992).
DR EMBL; S46270; AAP13884.1; -.
KW Transferase.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1516 MW; 113A46B08648F760 CRC64;

Query Match 28.8%; Score 30; DB 11; Length 13;
Best Local Similarity 55.6%; Pred. No. 4.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KVFPHGML 9
Db 3 KVFPHGVI 11

RESULT 2
Q9TWG5
ID Q9TWG5 PRELIMINARY; PRT; 19 AA.

OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;	
OC	campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;	
OC	Taraxacum.	
OX	NCBI_TaxID=154284;	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RP	TISSUE=Leaf;	
RA	Mes T.H.M.;	
RT	"Reconstruction of the evolution of trnF pseudogenes.";	
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AY015483; AAK21597.1; -.	
DR	GO: GO:0009507; C:chloroplast; IEA.	
KW	Chloroplast.	
FT	NON TER	1
SQ	SEQUENCE	16 AA; 1630 MW; 89810A9DAD549D92 CRC64;
Query Match	27.4%;	Score 28.5; DB 8; Length 16;
Best Local Similarity	56.2%;	Pred. No. 9.4e+02;
Matches	9; Conservative	0; Mismatches 4; Indels 3; Gaps 1
QY	3 NFFGHMLDLEEIAANS 18	
Db	1 NFF--LDLAAIEAPS 13	
RESULT 9		
Q95AY0	PRELIMINARY;	PRT; 16 AA.
ID	Q95AY0	
AC	Q95AY0;	
DT	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	
DR	PsbA protein (Fragment).	
DE	PSBA.	
GN	Artemisia rupestris.	
OS	Chloroplast	
OQ	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;	
OC	campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae;	
OC	Artemisia.	
OX	NCBI_TaxID=86317;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Peterson A., Peterson J.;	
RT	"Evolutionary aspects of Artemisia rupestris (L.).";	
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: AJ401617; CAC04434.2; -.	
DR	GO: GO:0009507; C:chloroplast; IEA.	
KW	Chloroplast.	
FT	NON TER	1
SQ	SEQUENCE	16 AA; 1672 MW; 93310A9DAD549D92 CRC64;
Query Match	27.4%;	Score 28.5; DB 8; Length 16;
Best Local Similarity	56.2%;	Pred. No. 9.4e+02;
Matches	9; Conservative	0; Mismatches 4; Indels 3; Gaps 1
QY	3 NFFGHMLDLEEIAANS 18	
Db	1 NFF--LDLAAIEAPS 13	
RESULT 10		
Q8SLF3	PRELIMINARY;	PRT; 16 AA.
ID	Q8SLF3	
AC	Q8SLF3;	
DT	01-JUN-2002 (TrEMBLrel. 21, Created)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)	
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	
DR	PsbA (Fragment).	
DE	PSBA.	
GN	Taraxacum (sect. Boreigena/Ceratophora) sp. t1148.	
OS	Chloroplast.	
OQ	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;
 OC Taraxacum.
 OX NCBI_TaxID=154243;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Mes T.H.M.;
 RT "Reconstruction of the evolution of trnF pseudogenes.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY015471; AAK21585.1; -.
 DR GO; GO:0009507; C:chloroplast; IEA.
 KW Chloroplast.
 FT NON TER
 SQ SEQUENCE 16 AA; 1630 MW; 89810A9DAD549D92 CRC64;
 Query Match 27.4%; Score 28.5; DB 8; Length 16;
 Best Local Similarity 56.2%; Pred. No. 9.4e+02;
 Matches 9; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
 QY 3 NFPHGMLDLEIEAANS 18
 ||| ||| ||| |||
 Db 1 NFP---LDLAAIEAPS 13

RESULT 11

Q8SLE1 ID Q8SLE1 PRELIMINARY; PRT; 17 AA.
 AC Q8SLE1; 17 AA.
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE PsBA (Fragment).
 GN PSBA.
 OS Taraxacum (sect. Scariosa) sp. 6057Hnew.
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;
 OC Taraxacum.
 OX NCBI_TaxID=154275;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Mes T.H.M.;
 RT "Reconstruction of the evolution of trnF pseudogenes.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY015484; AAK21598.1; -.
 DR GO; GO:0009507; C:chloroplast; IEA.
 KW Chloroplast.
 FT NON TER
 SQ SEQUENCE 17 AA; 1767 MW; F1868A9DAD549D92 CRC64;

Query Match 27.4%; Score 28.5; DB 8; Length 17;
 Best Local Similarity 56.2%; Pred. No. 1e+03;
 Matches 9; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
 QY 3 NFPHGMLDLEIEAANS 18
 ||| ||| ||| |||
 Db 2 NFP---LDLAAIEAPS 14

RESULT 12

Q8SLE8 ID Q8SLE8 PRELIMINARY; PRT; 17 AA.
 AC Q8SLE8; 17 AA.
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE PsBA (Fragment).
 GN PSBA.
 OS Taraxacum (sect. Tiberana) sp. JK4034.
 OC Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;
 OC Taraxacum.
 OX NCBI_TaxID=154278;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Mes T.H.M.;
 RT "Reconstruction of the evolution of trnF pseudogenes.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY015476; AAK21590.1; -.
 DR GO; GO:0009507; C:chloroplast; IEA.
 KW Chloroplast.
 FT NON TER
 SQ SEQUENCE 17 AA; 1767 MW; F1868A9DAD549D92 CRC64;
 Query Match 27.4%; Score 28.5; DB 8; Length 17;
 Best Local Similarity 56.2%; Pred. No. 1e+03;
 Matches 9; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
 QY 3 NFPHGMLDLEIEAANS 18
 ||| ||| ||| |||
 Db 2 NFP---LDLAAIEAPS 14

RESULT 13

Q8SLE0 ID Q8SLE0 PRELIMINARY; PRT; 17 AA.
 AC Q8SLE0; 17 AA.
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE PsBA (Fragment).
 GN PSBA.
 OS Taraxacum (sect. Naevosa) sp. 628Hnew.
 OC Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;
 OC Taraxacum.
 OX NCBI_TaxID=154265;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Mes T.H.M.;
 RT "Reconstruction of the evolution of trnF pseudogenes.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY015485; AAK21599.1; -.
 DR GO; GO:0009507; C:chloroplast; IEA.
 KW Chloroplast.
 FT NON TER
 SQ SEQUENCE 17 AA; 1767 MW; F1868A9DAD549D92 CRC64;

Query Match 27.4%; Score 28.5; DB 8; Length 17;
 Best Local Similarity 56.2%; Pred. No. 1e+03;
 Matches 9; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
 QY 3 NFPHGMLDLEIEAANS 18
 ||| ||| ||| |||
 Db 2 NFP---LDLAAIEAPS 14

RESULT 14

Q8SLG4 ID Q8SLG4 PRELIMINARY; PRT; 17 AA.
 AC Q8SLG4; 17 AA.
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE PsBA (Fragment).
 GN PSBA.
 OS Taraxacum (sect. Tiberana/Leucantha) sp. K3664.

Search completed: July 15, 2004, 14:50:59
Job time : 27.7333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:47:02 ; Search time 12.4 Seconds
(without alignments)
83.268 Million cell updates/sec

Title: US-09-171-432A-44
Perfect score: 104
Sequence: 1 KVNPHGMLEETAAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 204689

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	30.8	8	4	US-09-266-320D-4
2	32	30.8	16	6	5433940-10
3	31	29.8	20	3	US-08-256-747C-10
4	31	29.8	20	3	US-08-834-130A-10
5	31	29.8	20	4	US-09-387-671-46
6	31	29.8	21	4	US-09-010-317-20
7	31	29.8	21	4	US-09-010-317-34
8	30	28.8	26	4	US-09-227-357-333
9	29	27.9	15	3	US-08-595-945-10
10	29	27.9	17	3	US-08-454-098-26
11	29	27.9	20	1	US-08-487-001A-30
12	29	27.9	20	2	US-08-630-822A-30
13	29	27.9	20	2	US-09-005-069-30
14	29	27.9	20	4	US-09-004-730A-141
15	29	27.9	20	4	US-08-981-799A-141
16	29	27.9	26	2	US-08-875-062-5
17	28	26.9	10	3	US-09-043-930-14
18	28	26.9	19	2	US-08-564-972-72
19	28	26.9	21	3	US-08-787-091-7
20	28	26.9	22	6	5196404-10
21	28	26.9	22	6	5433940-9
22	28	26.9	25	1	US-07-987-286-20
23	28	26.9	25	2	US-08-614-626-20
24	28	26.9	28	3	US-09-040-485-5
25	28	26.9	30	2	US-08-459-568-49
26	28	26.9	30	2	US-08-399-411-49
27	28	26.9	30	3	US-08-516-859A-49

28	26.9	30	4	US-09-586-472-49	Sequence 49, Appl
29	26.9	30	4	US-09-528-706-49	Sequence 49, Appl
30	26.0	11	3	US-09-273-565-50	Sequence 50, Appl
31	26.0	11	4	US-09-565-538-50	Sequence 50, Appl
32	26.0	11	4	US-09-661-468-50	Sequence 50, Appl
33	26.0	11	4	US-09-976-165-50	Sequence 50, Appl
34	26.0	15	4	US-09-255-501-194	Sequence 194, App
35	26.0	15	4	US-09-255-501-195	Sequence 195, App
36	26.0	15	4	US-09-255-501-196	Sequence 196, App
37	26.0	15	4	US-09-387-671-38	Sequence 38, Appl
38	26.0	15	4	US-09-387-671-43	Sequence 43, Appl
39	26.0	16	3	US-08-855-531D-45	Sequence 45, Appl
40	26.0	16	3	US-08-855-526B-45	Sequence 45, Appl
41	26.0	16	4	US-09-387-671-37	Sequence 37, Appl
42	26.0	16	4	US-09-387-671-42	Sequence 42, Appl
43	26.0	17	4	US-09-387-671-36	Sequence 36, Appl
44	26.0	17	4	US-09-387-671-41	Sequence 41, Appl
45	26.0	18	4	US-09-387-671-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-266-320D-4
; Sequence 4, Application US/09266320D
; Patent No. 6664082
; GENERAL INFORMATION:
; APPLICANT: Hoshino, Tatsuo
; APPLICANT: Kon, Takahide
; APPLICANT: Shinjoh, Masako
; APPLICANT: Tazoe, Masaaki
; TITLE OF INVENTION: GENETICALLY ENGINEERED L-SORBOSE REDUCTASE-DEFICIENT MUTANTS
; FILE REFERENCE: 20079 US (C38435/109740)
; CURRENT APPLICATION NUMBER: US/09/266,320D
; CURRENT FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: EPO 98104546.1
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Gluconobacter suboxydans
US-09-266-320D-4

Query Match 30.8%; Score 32; DB 4; Length 8;
Best Local Similarity 71.4%; Pred. No. 36+05;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 FPHGMLE 10
DB 1 FPHGMLE 7

RESULT 2
5433940-10
; Patent No. 5433940
; APPLICANT: MARAGANORE, JOHN M.; FENTON II, JOHN M.; KLINE, TONI
; TITLE OF INVENTION: INHIBITORS OF THROMBIN
; NUMBER OF SEQUENCES: 25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,259
; FILING DATE: 17-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 548,388
; FILING DATE: 06-JUL-1989
; APPLICATION NUMBER: 395,482
; FILING DATE: 18-AUG-1989
; SEQ ID NO:10:
; LENGTH: 16
5433940-10

Query Match 30.8%; Score 32; DB 6; Length 16;
 Best Local Similarity 63.6%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 FPHGMLDEEI 14
 Db 1 FFXGNGDPEEI 11

RESULT 3
 US-08-256-747C-10
 ; Sequence 10, Application US/08256747C
 ; Patent No. 6037448
 ; GENERAL INFORMATION:
 ; APPLICANT: CHONG, Pele
 ; APPLICANT: GILLAM, Shirley
 ; APPLICANT: OU, Dawei
 ; APPLICANT: TINGLE, Aubrey
 ; TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR A RUBELLA VACCINE
 ; NUMBER OF SEQUENCES: 78
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sim & McBurney
 ; STREET: 6th Floor 701, 330 University Avenue
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5G 1R7
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/256,747C
 ; FILING DATE: 06-OCT-1994
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Stewart, Michael I.
 ; REGISTRATION NUMBER: 24,973
 ; REFERENCE/DOCKET NUMBER: 1038-370
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 595-1155
 ; TELEFAX: (416) 595-1163
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-256-747C-10

Query Match 29.8%; Score 31; DB 3; Length 20;
 Best Local Similarity 62.5%; Pred. No. 95;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NPHGMLD 10

Db 13 NTPHQLE 20

RESULT 4
 US-08-834-130A-10
 ; Sequence 10, Application US/08834130A
 ; Patent No. 6180758
 ; GENERAL INFORMATION:
 ; APPLICANT: CHONG, Pele
 ; APPLICANT: GILLAM, Shirley
 ; APPLICANT: OU, Dawei
 ; APPLICANT: TINGLE, Aubrey
 ; TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR A RUBELLA VACCINE
 ; NUMBER OF SEQUENCES: 75
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sim & McBurney
 ; STREET: 6th Floor 701, 330 University Avenue
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5G 1R7
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/256,747C
 ; FILING DATE: 06-OCT-1994
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Stewart, Michael I.
 ; REGISTRATION NUMBER: 24,973
 ; REFERENCE/DOCKET NUMBER: 1038-370
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 595-1155
 ; TELEFAX: (416) 595-1163
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-256-747C-10

Query Match 29.8%; Score 31; DB 3; Length 20;
 Best Local Similarity 62.5%; Pred. No. 95;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NPHGMLD 10

Db 13 NTPHQLE 20

RESULT 4
 US-08-834-130A-10
 ; Sequence 10, Application US/08834130A
 ; Patent No. 6180758
 ; GENERAL INFORMATION:
 ; APPLICANT: CHONG, Pele
 ; APPLICANT: GILLAM, Shirley
 ; APPLICANT: OU, Dawei
 ; APPLICANT: TINGLE, Aubrey
 ; TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR A RUBELLA VACCINE
 ; NUMBER OF SEQUENCES: 75
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sim & McBurney
 ; STREET: 6th Floor 701, 330 University Avenue
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5G 1R7
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/834,130A
 ; FILING DATE: 14-APR-1997
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: STEWART, Michael I.
 ; REGISTRATION NUMBER: 24,973
 ; REFERENCE/DOCKET NUMBER: 1038-686 MIS:jb
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 595-1155
 ; TELEFAX: (416) 595-1163
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-834-130A-10

Query Match 29.8%; Score 31; DB 3; Length 20;
 Best Local Similarity 62.5%; Pred. No. 95;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NPHGMLD 10

Db 13 NTPHQLE 20

RESULT 5
 US-09-387-671-46
 ; Sequence 46, Application US/09387671
 ; Patent No. 6673346
 ; GENERAL INFORMATION:
 ; APPLICANT: Ward, Peter A.
 ; APPLICANT: Huber-Lang, Markus
 ; APPLICANT: Sarma, Vidya
 ; APPLICANT: Czermak, Boris
 ; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
 ; FILE REFERENCE: UM-03783
 ; CURRENT APPLICATION NUMBER: US/09/387,671
 ; CURRENT FILING DATE: 1999-08-31
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 46
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; US-09-387-671-46

Query Match 29.8%; Score 31; DB 4; Length 20;
 Best Local Similarity 64.3%; Pred. No. 95;
 Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 8 MLD--LEETAANSK 19

Db 1 MLDKIEETAAKYK 14

RESULT 6
 US-09-010-317-20
 ; Sequence 20, Application US/09010317

```

; Patent No. 6685943
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen L.
; APPLICANT: Speciale, Pietro
; APPLICANT: Joh, Danny
; APPLICANT: McGavin, Martin J.
; TITLE OF INVENTION: FIBRONECTIN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,317
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,139
; FILING DATE: 21-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TAMK:189
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000
; TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-010-317-20

Query Match 29.8%; Score 31; DB 4; Length 21;
Best Local Similarity 55.6%; Pred. No. 16+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PHGMLDLEE 13
DB 8 PHNSVDLEE 16

RESULT 7
US-09-010-317-34
; Sequence 34, Application US/09010317
; Patent No. 6685943
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen L.
; APPLICANT: Speciale, Pietro
; APPLICANT: Joh, Danny
; APPLICANT: McGavin, Martin J.
; TITLE OF INVENTION: FIBRONECTIN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX

```

```

; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,317
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,139
; FILING DATE: 21-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TAMK:189
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000
; TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-010-317-34

Query Match 29.8%; Score 31; DB 4; Length 21;
Best Local Similarity 55.6%; Pred. No. 16+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 PHGMLDLEE 13
DB 8 PHNSVDLEE 16

RESULT 8
US-09-227-357-333
; Sequence 333, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010PI
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920

```

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 333
LENGTH: 26
TYPE: PRT
ORGANISM: Homo sapiens
US-09-227-357-333

Query Match 28.8%; Score 30; DB 4; Length 26;
Best Local Similarity 53.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KVNPFHGMLEEE 13
Db 6 KVDLEHYMLVLKE 18
||: |||||:|
||: |||||:|

RESULT 9
US-08-595-945-10
Sequence 10, Application US/08595945
Patent No. 6143509
GENERAL INFORMATION:
APPLICANT: DOWELL, BARRY L.
APPLICANT: BRIDON, DOMINIQUE P.
APPLICANT: QIU, XIAOXING
APPLICANT: LILJA, HANS
APPLICANT: PIIRONEN, TIMO P.
APPLICANT: VIHINEN, MAUNO A.
APPLICANT: PETTERSSON, KIM S. I.
TITLE OF INVENTION: PROSTATE SPECIFIC ANTIGEN PEPTIDES
AND USES THEREOF
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: 100 ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/595,945
APPLICATION NUMBER: US/08/595,945
FILING DATE: 06-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BECKER, CHERYL L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 5875.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-935-1729
TELEFAX: 847-938-2623
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-595-945-10

Query Match 27.9%; Score 29; DB 3; Length 15;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NFRHGMLDL 11
Db 3 SFPHPLYDM 11
:|:|:|:|:
:|:|:|:|:

RESULT 10
US-08-454-098-26
Sequence 26, Application US/08454098
Patent No. 6103521
GENERAL INFORMATION:
APPLICANT: CAPON, DANIEL J
APPLICANT: SMITH, DOUGLAS H
APPLICANT: TIAN, HUAN
APPLICANT: WINSLOW, GENINE A
APPLICANT: SIEKEVITZ, MIRIAM
TITLE OF INVENTION: MULTISPECIFIC CHIMERIC RECEPTORS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CELL GENESYS, INC.
STREET: 322 LAKESIDE DRIVE
CITY: FOSTER CITY
STATE: CALIFORNIA
COUNTRY: US
ZIP: 94404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,098
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/384,033
FILING DATE:
ATTORNEY/AGENT INFORMATION:

CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/005,069
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/630,822
 FILING DATE: 11-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: CONNELL, GARY J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2618-17-C3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: Xaa = any amino acid
 LOCATION: 15
 US-09-005-069-30

Query Match 27.9%; Score 29; DB 2; Length 20;
 Best Local Similarity 38.5%; Pred. No. 2.1e+02;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 KVNFPFGHMLDLEE 13
 Db 1 EVSIPSGKLTIED 13

RESULT 14
 US-09-004-730A-141
 ; Sequence 141, Application US/09004730A
 ; Patent No. 6485368
 ; GENERAL INFORMATION:
 ; APPLICANT: Weber, Eric
 ; APPLICANT: Wu Hunter, Shirley
 ; APPLICANT: Sim, Gek-Kee
 ; APPLICANT: Frank, Glenn
 ; APPLICANT: Wallenfels, Lynda
 ; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO COLLECT SUCH
 ; FILE REFERENCE: 2618-17-C5-PUS-1
 ; CURRENT APPLICATION NUMBER: US/09/004,730A
 ; CURRENT FILING DATE: 1998-01-08
 ; PRIOR APPLICATION NUMBER: PCT/97US/18669
 ; PRIOR FILING DATE: 1997-10-15
 ; NUMBER OF SEQ ID NOS: 150
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 141
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Ctenocephalides felis
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (15)..()
 ; OTHER INFORMATION: Xaa = any amino acid
 US-09-004-730A-141

Query Match 27.9%; Score 29; DB 4; Length 20;
 Best Local Similarity 38.5%; Pred. No. 2.1e+02;

Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 1 KVNFPFGHMLDLEE 13
 Db 1 EVSIPSGKLTIED 13

RESULT 15
 US-08-981-799A-141
 ; Sequence 141, Application US/08981799A
 ; Patent No. 6576238
 ; GENERAL INFORMATION:
 ; APPLICANT: Weber, Eric
 ; APPLICANT: Wu Hunter, Shirley
 ; APPLICANT: Sim, Gek-Kee
 ; APPLICANT: Frank, Glenn
 ; APPLICANT: Wallenfels, Lynda
 ; TITLE OF INVENTION: "NOVEL ECTOPARASITE SALIVA PROTEINS AND APPARATUS TO COLLECT SUCH
 ; FILE REFERENCE: 2618-17-C5-PUS
 ; CURRENT APPLICATION NUMBER: US/08/981,799A
 ; CURRENT FILING DATE: 1998-08-27
 ; PRIOR APPLICATION NUMBER: PCT/97/18669
 ; PRIOR FILING DATE: 1997-10-15
 ; NUMBER OF SEQ ID NOS: 150
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 141
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Ctenocephalides felis
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (15)..()
 ; OTHER INFORMATION: Xaa = any amino acid
 US-08-981-799A-141

Query Match 27.9%; Score 29; DB 4; Length 20;
 Best Local Similarity 38.5%; Pred. No. 2.1e+02;
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 KVNFPFGHMLDLEE 13
 Db 1 EVSIPSGKLTIED 13

Search completed: July 15, 2004, 15:13:21
 Job time : 13.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:11:48 ; Search time 32.5333 Seconds
(without alignments)
192.148 Million cell updates/sec

Title: US-09-171-432A-44
Perfect score: 104
Sequence: 1 KUNFPFGMDLEETIAANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 288454

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	20	10	US-09-171-432A-44
2	53	51.0	20	10	US-09-171-432A-45
3	36	34.6	25	16	US-10-666-480-120
4	34	32.7	20	16	US-10-666-480-19
5	34	32.7	21	16	US-10-666-480-48
6	33	31.7	14	10	US-09-880-748-2513
7	33	31.7	14	12	US-10-293-418-2513
8	33	31.7	16	14	US-10-225-567A-1118
9	32	30.8	14	10	US-09-880-748-2152
10	32	30.8	14	12	US-10-293-418-2152
11	32	30.8	19	16	US-10-181-183A-5
12	32	30.8	23	10	US-09-927-734C-5
13	31	29.8	20	9	US-09-878-603-46
14	31	29.8	27	12	US-10-424-599-202685
15	31	29.8	30	12	US-10-296-734-1382

16	31	29.8	30	12	US-10-296-734-1384
17	30.5	29.3	22	9	US-09-739-907-167
18	30.5	29.3	22	11	US-09-938-671-167
19	30	28.8	12	10	US-09-954-385-284
20	30	28.8	14	10	US-09-880-748-2473
21	30	28.8	14	12	US-10-293-418-2473
22	30	28.8	21	14	US-10-062-710-176
23	30	28.8	23	12	US-10-424-599-173402
24	30	28.8	26	10	US-09-983-802-333
25	30	28.8	26	12	US-09-973-278-358
26	30	28.8	26	12	US-09-984-490-333
27	29	27.9	13	12	US-09-988-493-270
28	29	27.9	13	12	US-10-014-340-565
29	29	27.9	14	10	US-09-880-748-2369
30	29	27.9	14	10	US-09-880-748-2400
31	29	27.9	14	10	US-09-880-748-2402
32	29	27.9	14	10	US-09-880-748-2404
33	29	27.9	14	10	US-09-880-748-2444
34	29	27.9	14	10	US-09-880-748-2475
35	29	27.9	14	10	US-09-880-748-2538
36	29	27.9	14	10	US-09-880-748-2631
37	29	27.9	14	12	US-10-293-418-2369
38	29	27.9	14	12	US-10-293-418-2400
39	29	27.9	14	12	US-10-293-418-2402
40	29	27.9	14	12	US-10-293-418-2404
41	29	27.9	14	12	US-10-293-418-2444
42	29	27.9	14	12	US-10-293-418-2475
43	29	27.9	14	12	US-10-293-418-2538
44	29	27.9	14	12	US-10-293-418-2631
45	29	27.9	14	14	US-10-239-313A-400

ALIGNMENTS

RESULT 1
US-09-171-432A-44
; Sequence 44, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Khudiyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label= YK-1327
US-09-171-432A-44

Query Match 100.0%; Score 104; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNPHGMLDLEETAANSKD 20
Db 1 KNPHGMLDLEETAANSKD 20
|||||

RESULT 2

US-09-171-432A-45
; Sequence 45, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499
INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label= YK-1328
US-09-171-432A-45

Query Match 51.0%; Score 53; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 DLEETAANSKD 20
Db 1 DLEETAANSKD 11
|||||

RESULT 3

US-10-666-480-120
; Sequence 120, Application US/106666480
; Publication No. US20040121959A1
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C
; APPLICANT: Wild, Kenneth D
; APPLICANT: Sitney, Karen C
; APPLICANT: Min, Hosung
; APPLICANT: Kimmel, Bruce
; TITLE OF INVENTION: Peptides and Related Molecules That Modulate Nerve Growth Factor
; FILE REFERENCE: A-827US
; CURRENT APPLICATION NUMBER: US/10/666,480
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/412,524
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Therapeutically active peptide of randomly generated, non-natural
; OTHER INFORMATION: ly occurring sequence
US-10-666-480-120

Query Match 34.6%; Score 36; DB 16; Length 25;
Best Local Similarity 63.6%; Pred. No. 69;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NPHGMLDLEE 13
Db 13 NPHGMLDLEE 23
|||||

RESULT 4

US-10-666-480-19
; Sequence 19, Application US/106666480
; Publication No. US20040121959A1
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C
; APPLICANT: Wild, Kenneth D
; APPLICANT: Sitney, Karen C
; APPLICANT: Min, Hosung
; APPLICANT: Kimmel, Bruce
; TITLE OF INVENTION: Peptides and Related Molecules That Modulate Nerve Growth Factor
; FILE REFERENCE: A-827US
; CURRENT APPLICATION NUMBER: US/10/666,480
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/412,524
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Therapeutically active peptide of randomly generated, non-natural
; OTHER INFORMATION: ly occurring sequence
US-10-666-480-19

Query Match 32.7%; Score 34; DB 16; Length 20;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NPHGMLDLEE 11


```
Db      12 NFPPGVLEL 20      ||| |||:|
                                           ||| |||
                                           7 FPHAPLDL 14

RESULT 5
US-10-666-480-48
; Sequence 48, Application US/10666480
; Publication No. US20040121959A1
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C
; APPLICANT: Wild, Kenneth D
; APPLICANT: Sitney, Karen C
; APPLICANT: Min, Hosung
; APPLICANT: Kimmel, Bruce
; TITLE OF INVENTION: Peptides and Related Molecules That Modulate Nerve Growth Factor
; FILE REFERENCE: A-827US
; CURRENT APPLICATION NUMBER: US/10/666,480
; CURRENT FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: 60/412,524
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Therapeutically active peptide of randomly generated, non-natural
; OTHER INFORMATION: ly occurring sequence
US-10-666-480-48

Query Match      32.7%; Score 34; DB 16; Length 21;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 NFPHGMLDL 11      ||| |||:|
                                           ||| |||
                                           7 FPHAPLDL 14

Db      13 NFPPGVLEL 21

RESULT 6
US-09-880-748-2513
; Sequence 2513, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2513
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-2513

Query Match      31.7%; Score 33; DB 10; Length 14;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 FPHGMLDL 11

RESULT 7
US-10-293-418-2513
; Sequence 2513, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 2513
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-2513

Query Match      31.7%; Score 33; DB 12; Length 14;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 FPHGMLDL 11
                                           ||| |||
                                           7 FPHAPLDL 14

Db      7 FPHAPLDL 14

RESULT 8
US-10-225-567A-1118
; Sequence 1118, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1118
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-1118

Query Match      31.7%; Score 33; DB 14; Length 16;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

QY 5 PHGMLDLLE 13
||:||||
Db 5 PHSVIDYEE 13

RESULT 9

US-09-880-748-2152
; Sequence 2152, Application US/09880748
; Publication No. US2003005937A1
; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

; FILE REFERENCE: PF523

; CURRENT APPLICATION NUMBER: US/09/880,748

; CURRENT FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 3239

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2152

; LENGTH: 14

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-880-748-2152

Query Match 30.8%; Score 32; DB 10; Length 14;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PHGMLDL 11
||:||||
Db 7 PPHSLDL 14

RESULT 10

US-10-293-418-2152
; Sequence 2152, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys

; FILE REFERENCE: PF523P2

; CURRENT APPLICATION NUMBER: US/10/293,418

; CURRENT FILING DATE: 2002-11-27

; PRIOR APPLICATION NUMBER: 60/331,469

; PRIOR FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: 60/340,817

; PRIOR FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 03/880,748

; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: 60/293,499

; PRIOR FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: 60/277,379

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/240,816

; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/212,210

; PRIOR FILING DATE: 2000-06-16

; NUMBER OF SEQ ID NOS: 3247

; SEQ ID NO 2152

; LENGTH: 14

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-293-418-2152

Query Match 30.8%; Score 32; DB 12; Length 14;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPHGMLDL 11
||:||||
Db 7 PPHSLDL 14

RESULT 11

US-10-181-183A-5

; Sequence 5, Application US/10181183A

; Publication No. US20040121431A1

; GENERAL INFORMATION:

; APPLICANT: KUBOTA, Michio

; APPLICANT: MARUTA, Kazuhiko

; APPLICANT: YAMAMOTO, Takuo

; APPLICANT: FUKUDA, Shigeharu

; TITLE OF INVENTION: Polypeptide having alpha-isomaltosyl-transferring enzymatic acti

; FILE REFERENCE: KUBOTA=10

; CURRENT APPLICATION NUMBER: US/10/181,183A

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 19

; SEQ ID NO 5

; LENGTH: 19

; TYPE: PRT

; ORGANISM: Bacillus globisporus

US-10-181-183A-5

Query Match 30.8%; Score 32; DB 16; Length 19;
Best Local Similarity 58.3%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 PHGMLDLLEIAA 16
||:||||
Db 8 PYGIDDLVEIOA 19

RESULT 12

US-09-927-734C-5

; Sequence 5, Application US/09927734C

; Publication No. US20030194758A1

; GENERAL INFORMATION:

; APPLICANT: Selitrennikoff, Claude

; APPLICANT: Nakata, Mitsunori

; TITLE OF INVENTION: Method for the Identification of Fungal Glucose Utilization

; FILE REFERENCE: MYCOLOGX-06279

; CURRENT APPLICATION NUMBER: US/09/927,734C

; CURRENT FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 5

; LENGTH: 23

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-927-734C-5

Query Match 30.8%; Score 32; DB 10; Length 23;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 LEEIAANSK 19
||:||||
Db 1 IEDIAVNSK 9

RESULT 13

US-09-878-603-46

; Sequence 46, Application US/09878603

; Patent No. US20020165138A1

```
; GENERAL INFORMATION:
; APPLICANT: Ward, Peter A.
; APPLICANT: Huber-Lang, Markus
; APPLICANT: Sarma, Vidya
; APPLICANT: Czermak, Boris
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Sepsis
; FILE REFERENCE: UM-03783
; CURRENT APPLICATION NUMBER: US/09/878,603
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/387,671
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 46
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-09-878-603-46

Query Match      29.8%; Score 31; DB 9; Length 20;
Best Local Similarity 64.3%; Pred. No. 3.8e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY      8 MLD--LEETAAASK 19
      ||| :||||| |
Db      1 MLDKTEETAAKYK 14

RESULT 14
US-10-424-599-202685
; Sequence 202685, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovacic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 202685
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_25049C.1.pep
; US-10-424-599-202685

Query Match      29.8%; Score 31; DB 12; Length 27;
Best Local Similarity 42.9%; Pred. No. 5.3e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY      1 KVNPPHGMLEEL 14
      |||| |::: |
Db      3 KVNPF--GLIDISHV 14

RESULT 15
US-10-296-734-1382
; Sequence 1382, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
```

```
; PRIOR APPLICATION NUMBER: AU PQ7761/00
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1382
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: PRAME segment 30
; US-10-296-734-1382

Query Match      29.8%; Score 31; DB 12; Length 30;
Best Local Similarity 60.0%; Pred. No. 6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      6 HGMLEEL 15
      ||| :||| |
Db      19 HGTTLHLERLA 28

Search completed: July 15, 2004, 15:22:56
Job time : 33.5333 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:36:31 ; Search time 40.1333 Seconds
(without alignments)
140.804 Million cell updates/sec

Title: US-09-171-432A-45
Perfect score: 101
Sequence: 1 DLEETAAANSKDFPNMSETDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 581562

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_29Jan04:*
- 1: Geneseq1980s:*
 - 2: Geneseq1990s:*
 - 3: Geneseq2000s:*
 - 4: Geneseq2001s:*
 - 5: Geneseq2002s:*
 - 6: Geneseq2003as:*
 - 7: Geneseq2003bs:*
 - 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	101	100.0	20	2	AAW42928	AAW42928 Immunogen
2	101	100.0	21	4	AAB69445	AAB69445 Synthetic
3	53	52.5	20	2	AAW42927	AAW42927 Immunogen
4	53	52.5	21	4	AAB69444	AAB69444 Synthetic
5	38	37.6	21	2	AAR06538	AAR06538 Human B-C
6	38	37.6	26	2	AAR06536	AAR06536 Human B-C
7	38	37.6	30	2	AAR06535	AAR06535 Human B-C
8	35	34.7	10	3	AAB09935	AAB09935 Human can
9	35	34.7	10	5	ABG79071	ABG79071 Human ART
10	35	34.7	10	6	ABR84379	ABR84379 Human ART
11	35	34.7	10	7	ADC17704	ADC17704 ART4 prot
12	35	34.7	22	3	AAB51604	AAB51604 Yada homo
13	35	34.7	24	3	AAB51532	AAB51532 Proteobac
14	35	34.7	29	2	AAR57372	AAR57372 Ala-103 L
15	33	32.7	12	6	ADA08124	ADA08124 Alpha ID
16	33	32.7	15	2	AAW95342	AAW95342 Phospholi
17	33	32.7	20	2	AAR37216	AAR37216 The psi-7
18	33	32.7	20	7	ADE15764	ADE15764 E. coli t
19	33	32.7	25	5	ABG62250	ABG62250 Eubacteri
20	33	32.7	27	4	AAB91127	AAB91127 Parathyro
21	33	32.7	29	6	ABU71001	ABU71001 Human adi
22	32	31.7	15	5	AAE16442	AAE16442 Oestrogen
23	32	31.7	15	5	AAU86354	AAU86354 Oestrogen
24	32	31.7	15	5	AAU86406	AAU86406 Oestrogen
25	32	31.7	15	5	AAU86406	AAU86406 Oestrogen

ALIGNMENTS

RESULT 1									
AAW42928									
ID	AAW42928	standard;	peptide;	20	AA.				
XX	AC	AAW42928;							
XX	AC	AAW42928;							
DT	28-APR-1998	(first entry)							
XX	DE	Immunogenic Hepatitis A virus peptide YK-1328.							
XX	DE	Immunogenic Hepatitis A virus peptide YK-1328.							
KW	Immunogenic peptide;	immunogenic epitope;	P2A protein;	immune response;					
KW	antibody.								
XX	OS	Synthetic.							
OS	Hepatitis A virus.								
XX	PN	WO9740147-A1.							
XX	PD	30-OCT-1997.							
XX	PF	18-APR-1997;	97WO-US006891.						
XX	PR	19-APR-1996;	96US-0015644P.						
XX	PA	(USSH)	US DEPT HEALTH & HUMAN SERVICES.						
XX	FI	Fields HA,	Khudyakov YE;						
XX	DR	WPI;	1997-535831/49.						
XX	PT	Immunogenic Hepatitis A Virus (HAV) peptide(s) -	used to induce an immune						
XX	PT	response to HAV in a mammal or to detect the presence of antibodies	against HAV in a mammal.						
XX	PS	Claim 18;	Page 112; 140pp; English.						
XX	CC	Peptides AAW42922-30 are immunogenic peptides corresponding to	immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are						
XX	CC	substantially similar to a portion of the amino acid sequence of the P2A	protein of HAV corresponding to amino acids 931-950. The present peptide						
XX	CC	is derived from amino acids 931-950, and has a reactivity of 12.5% with	acute sera. Compositions containing the peptides can be used to induce an						
XX	CC	immune response to HAV in a mammal. The peptides can also be used to	detect the presence of antibodies against HAV in mammalian serum. The						
XX	CC	peptides can also be used to make an antibody against HAV by	administering the peptide to a mammal						
XX	XX	Sequence 20 AA;							

XX WO200105824-A2.
 PN 25-JAN-2001.
 PD 14-JUL-2000; 2000WO-US019267.
 XX 15-JUL-1999; 99US-0144412P.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Fields HA, Khudyakov YE;
 PI WPI; 2001-112681/12.
 DR Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 PT Claim 13; Page 96; 130pp; English.
 XX The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum.
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IGM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IGM antibody reactivity
 XX Sequence 21 AA;
 SQ
 Query Match 52.5%; Score 53; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 DLEETAAANSKD 11
 Db 10 DLEETAAANSKD 20
 RESULT 5
 ID AAR06538 standard; protein; 21 AA.
 XX AAR06538;
 AC AAR06538;
 XX 25-MAR-2003 (revised)
 DT 04-JAN-1991 (first entry)
 PT Human B-cell stimulator 2 (BSF2) antagonist.
 DE Immunodeficiency; myeloma; chronic articular rheumatism; endotoxin shock.
 XX Homo sapiens.
 XX JP02188600-A.
 XX 24-JUL-1990.
 PD 17-JAN-1989; 89JP-00007944.
 PF 17-JAN-1989; 89JP-00007944.
 XX (CHUS) CHUGAI PHARM CO LTD.
 PA WPI; 1990-266224/35.
 DR Polypeptide with antagonist against human B cell stimulator 2 - comprises
 PT human B cell stimulator with deficiency of 2 or more aminoacid(s) at N-
 PT and/or C-terminal myocardial infarction.
 XX Claim 9; Page 968; 6pp; Japanese.
 PS BSF2 antagonists are useful in diagnosis and treatment of immuno-
 CC deficiencies eg. myeloma, chronic articular rheumatism and endotoxin
 CC shock. Antagonist comprises BSF2 peptide (pref. AAs 20-40) with
 CC substitutions of two or more C- and/or N-terminal AAs. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX Sequence 26 AA;
 SQ
 Query Match 37.6%; Score 38; DB 2; Length 26;
 Best Local Similarity 41.2%; Pred. No. 51;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 OY 3 BEIAANSKDFPNMSETD 19
 Db 7 EALAENNLNLPKMAEKD 23

DR WPI; 1990-266224/35.
 XX Polypeptide with antagonist against human B cell stimulator 2 - comprises
 PT human B cell stimulator with deficiency of 2 or more aminoacid(s) at N-
 PT and/or C-terminal myocardial infarction.
 XX Claim 11; Page 968; 6pp; Japanese.
 XX BSF2 antagonists are useful in diagnosis and treatment of immuno-
 CC deficiencies eg. myeloma, chronic articular rheumatism and endotoxin
 CC shock. Antagonist comprises BSF2 peptide (pref. AAs 20-40) with
 CC substitutions of two or more C- and/or N-terminal AAs. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX Sequence 21 AA;
 SQ
 Query Match 37.6%; Score 38; DB 2; Length 21;
 Best Local Similarity 41.2%; Pred. No. 40;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 OY 3 BEIAANSKDFPNMSETD 19
 Db 5 EALAENNLNLPKMAEKD 21

RESULT 6
 ID AAR06536 standard; protein; 26 AA.
 XX AAR06536;
 AC AAR06536;
 XX 25-MAR-2003 (revised)
 DT 04-JAN-1991 (first entry)
 PT Human B-cell stimulator 2 (BSF2) antagonist.
 DE Immunodeficiency; myeloma; chronic articular rheumatism; endotoxin shock.
 XX Homo sapiens.
 XX JP02188600-A.
 XX 24-JUL-1990.
 PD 17-JAN-1989; 89JP-00007944.
 PF 17-JAN-1989; 89JP-00007944.
 XX (CHUS) CHUGAI PHARM CO LTD.
 PA WPI; 1990-266224/35.
 DR Polypeptide with antagonist against human B cell stimulator 2 - comprises
 PT human B cell stimulator with deficiency of 2 or more aminoacid(s) at N-
 PT and/or C-terminal myocardial infarction.
 XX Claim 9; Page 968; 6pp; Japanese.
 PS BSF2 antagonists are useful in diagnosis and treatment of immuno-
 CC deficiencies eg. myeloma, chronic articular rheumatism and endotoxin
 CC shock. Antagonist comprises BSF2 peptide (pref. AAs 20-40) with
 CC substitutions of two or more C- and/or N-terminal AAs. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX Sequence 26 AA;
 SQ
 Query Match 37.6%; Score 38; DB 2; Length 26;
 Best Local Similarity 41.2%; Pred. No. 51;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 OY 3 BEIAANSKDFPNMSETD 19
 Db 7 EALAENNLNLPKMAEKD 23

RESULT 7

AAR06535
ID AAR06535 standard; protein; 30 AA.

XX
AC AAR06535;

XX
DT 25-MAR-2003 (revised)

XX
DT 04-JAN-1991 (first entry)

XX
XX Human B-cell simulator 2 (BSP2) antagonist.

XX
KW Immunodeficiency; myeloma; chronic articular rheumatism; endotoxin shock.

XX
OS Homo sapiens.

XX
PN JP02188600-A.

XX
PD 24-JUL-1990.

XX
PF 17-JAN-1989; 89JP-00007944.

XX
PR 17-JAN-1989; 89JP-00007944.

XX
PA (CHUS) CHUGAI PHARM CO LTD.

XX
DR WPI; 1990-265224/35.

XX
PT Polypeptide with antagonist against human B cell simulator 2 - comprises human B cell stimulator with deficiency of 2 or more aminoacid(s) at N- and/or C-terminal myocardial infarction.

XX
PS Claim 8; Page 968; 6pp; Japanese.

XX
CC BSP2 antagonists are useful in diagnosis and treatment of immuno-deficiencies eg. myeloma, chronic articular rheumatism and endotoxin shock. Antagonist comprises BSP2 peptide (pref. AAs 20-40) with substitutions of two or more C- and/or N-terminal AAs. (Updated on 25-MAR-2003 to correct PA field.)

XX
SQ Sequence 30 AA;

Query Match 37.6%; Score 38; DB 2; Length 30;

Best Local Similarity 41.2%; Pred. No. 60; Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 3 EEIAANGKDFPNMSETD 19

Db 11 EALAEENLNLPKMAEKD 27

RESULT 8

AAB09935
ID AAB09935 standard; peptide; 10 AA.

XX
AC AAB09935;

XX
DT 19-OCT-2000 (first entry)

XX
XX Human cancer regression antigen protein fragment P6.

XX
KW Cancer regression antigen; human; cytostatic; tumor antigen protein;

XX
KW MHC class I antigen.

XX
OS Homo sapiens.

XX
PN JP2000116383-A.

XX
PD 25-APR-2000.

XX
PF 14-OCT-1998; 98JP-00291702.

XX

PR 14-OCT-1998; 98JP-00291702.

XX
PA (ITOY/) ITO Y.

XX
DR WPI; 2000-369405/32.

XX
PT Human cancer regression antigen protein, useful for the treatment of tumors.

XX
PS Disclosure; Page 6; 15pp; Japanese.

XX
CC This invention describes a novel polynucleotide (N1) selected from a polynucleotide encoding a 412 amino acid protein (P1) which has cytostatic activity. INDEPENDENT CLAIMS are included for the following: (1) a tumor antigen protein encoded by the above polynucleotide molecule; (2) an oligonucleotide molecule consisting of part of the above polynucleotide molecule and encoding a tumor antigen peptide recognized by T cell by combining with the MHC class I antigen; (3) an oligonucleotide molecule consisting of a coding sequence of a polynucleotide molecule encoding P1 or a sequence complementary to the base sequence in its 5' noncoding sequence or its chemically modified product; (4) a tumor antigen peptide encoded by the above oligonucleotide molecule or its derivative; (5) a drug containing the above tumor antigen protein or the above tumor antigen peptide or its derivative; and (6) an antibody against the above tumor antigen protein or the above tumor antigen peptide or its derivative. This sequence represents a fragment of the human cancer regression antigen protein described in the invention

XX
SQ Sequence 10 AA;

Query Match 34.7%; Score 35; DB 3; Length 10;

Best Local Similarity 50.0%; Pred. No. 53; Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 11 DFPNMSSTD 20

Db 1 DYPSSLATDI 10

RESULT 9

ABG79071
ID ABG79071 standard; peptide; 10 AA.

XX
AC ABG79071;

XX
DT 15-NOV-2002 (first entry)

XX
XX Human ART-4 class I HLA widely expressed antigen peptide #2.

XX
KW Cell penetrating peptide; cancer; tumour; melanoma; thymoma; antigen; lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia; Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer; kidney cancer; adenocarcinoma; breast cancer; prostate cancer; ovarian cancer; pancreatic cancer; epitope; vaccine; dendritic cell; tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA; cytostatic; human.

XX
OS Homo sapiens.

XX
PN WO200264057-A2.

XX
PD 22-AUG-2002.

XX
PF 15-FEB-2002; 2002WO-US005212.

XX
PR 15-FEB-2001; 2001US-0268687P.

XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX
PI Wang R;

XX
DR WPI; 2002-627577/67.

XX

PT Novel composition for treating a disease in an animal, comprises an
PT immune effector cell and cell penetrating peptide associated with an
PT antigen or antibody.
XX
PS Disclosure; Page 17; 6lpp; English.
XX
CC The invention relates to a composition (I) comprising an immune effector
CC cell and a cell penetrating peptide (CPP) associated with an antigen or
CC antibody. Also included are (1) a vaccine comprising (1), CPP associated
CC with an antigen, and a pharmaceutically acceptable carrier and (2)
CC preparing a composition for a disease, by providing (1) and CPP
CC associated with an antigen for a disease, and introducing the antigen-
CC associated CPP to (1), where antigen enters into the cell. The antigens
CC are, for example, tumour antigen derived epitopes recognised by tumour
CC infiltrating lymphocytes (TIL) of HLA (human leukocyte antigen) class I
CC or II. The composition is useful for enhancing immunity in an animal to a
CC disease, by administering a mature dendritic cell comprising CPP
CC associated with an antigen to disease, to the animal, such that following
CC the administration, animal is protected from disease, where the animal
CC comprises both CD4+ and CD8+ T cells. It is also useful for treating a
CC disease (e.g. cancer, tumour, melanoma, thymoma, lymphoma, sarcoma, lung
CC cancer, non-Hodgkin's lymphoma, leukaemia, Hodgkin's lymphoma, uterine
CC cancer, cervical cancer, bladder cancer, kidney cancer, adenocarcinoma,
CC breast cancer, prostate cancer, ovarian cancer and pancreatic cancer).
CC The animal is further subjected to a cancer treatment including surgery,
CC radiation, chemotherapy or gene therapy. The administration of (1),
CC preferably dendritic cell is prior to, subsequent to or concurrent with,
CC the cancer treatment. The present sequence is a tumour antigen derived
CC epitope for inclusion in the composition of the invention
XX
SQ Sequence 10 AA;
Query Match 34.7%; Score 35; DB 5; Length 10;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 11 DFPNMSETDL 20
Db 1 DYPSLSATDI 10
|:|:|:|:
|:|:|:|:
RESULT 10
ABR84379
ID ABR84379 standard; peptide; 10 AA.
AC ABR84379;
XX
DT 06-NOV-2003 (first entry)
DE Human ART4 HLA-A24 epitope, SEQ ID NO:29.
XX
XX Antigen specific T-cell; detection; diagnosis; cancer specific T-cell;
KW cancer; tumour; cervical cancer; prostate cancer; cellular immunity;
KW immune therapy; cytostatic; immunostimulant; vaccine; antigenic peptide;
KW human; human leukocyte antigen; HLA-A24 epitope.
XX
OS Homo sapiens.
XX
XX JP2002365286-A.
EN
XX
PD 18-DEC-2002.
XX
XX 18-SEP-2001; 2001JP-00283413.
PF
XX
PR 13-NOV-2000; 2000JP-00345094.
XX
XX (ITOY/) ITO Y.
FA
XX
DR WPI; 2003-508315/48.
XX
XX A detection method of antigen specific T-cells, comprises the use of
PT plural antigenic peptides, useful in semi-quantitative determination of
PT cancer specific T-cell frequencies and for monitoring cellular immunity.
PT

XX
PS Example 8; Page 10; 18pp; Japanese.
XX
CC The invention relates to a method for the detection of antigen specific T
CC -cells in a blood sample involving the use of a plurality of antigenic
CC peptides. The method comprises sampling of peripheral blood monocytes;
CC stimulation of the collected peripheral blood monocytes with antigens
CC without direct use of antigen presenting cells; and detection of T-cells
CC specific to the antigen in the stimulated monocytes. The method is
CC particularly used for the detection of cancer as it can be used in semi-
CC quantitative determination of cancer specific T-cells. It can also be
CC used for cancer vaccine therapy for patients with cervical or prostate
CC cancer. The method can additionally be used to monitor of cellular
CC immunity and cancer immune therapy by detection of specific T-cell
CC frequencies. Sequences ABR84367-ABR84380 represent HLA-A24 (human
CC leukocyte antigen) peptides of human origin used in an example from the
CC invention
XX
SQ Sequence 10 AA;
Query Match 34.7%; Score 35; DB 6; Length 10;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 11 DFPNMSETDL 20
Db 1 DYPSLSATDI 10
|:|:|:|:
|:|:|:|:
RESULT 11
ADCL7704
ID ADCL7704 standard; peptide; 10 AA.
XX
AC ADCL7704;
XX
DT 18-DEC-2003 (first entry)
DE ART4 protein amino acids 75-84.
XX
XX cytostatic; antiallergic; vaccine; allergic reaction suppressor; antigen;
KW allergy; epitope; cancer; immunoglobulin E antibody; desensitization;
KW ART4.
XX
OS Synthetic.
XX
XX WO2003020306-A1.
PN
XX
PD 13-MAR-2003.
XX
XX 28-AUG-2002; 2002WO-JP008641.
PF
XX
XX 29-AUG-2001; 2001JP-00260046.
PR
XX
XX (ITOH/) ITOH K.
PA
XX
PI Itoh K, Yamada A;
XX
XX WPI; 2003-300831/29.
DR
XX
XX Desensitizers or allergic reaction suppressors containing peptides
PT originating from the same antigenic substances as reaction-inducing
PT substances, useful for preventing or treating type I allergic diseases
PT and in cancer vaccines.
XX
XX Claim 9; SEQ ID NO 4; 49pp; Japanese.
XX
XX The invention relates to allergic reaction suppressors, comprising a
CC peptide which originates from the same antigenic substance as the
CC antigenic substance inducing the (peptide-originated) allergic reaction,
CC and containing an epitope different from the epitope participating in the
CC induction of the allergic reaction and yet does not induce the allergic
CC reaction. The desensitizers or allergic reaction suppressors are useful
CC for treating and preventing type I allergic diseases and in cancer

CC vaccines for preventing or treating cancer. The allergic reaction
 CC suppressors are also useful for suppressing an allergic reaction or
 CC reducing immunoglobulin E antibody production, and for desensitization to
 CC antigens. This sequence represents a peptide used in the invention and
 CC corresponds to amino acids 75-84 of the ART4 protein.
 XX
 SQ Sequence 10 AA;

Query Match 34.7%; Score 35; DB 7; Length 10;
 Best Local Similarity 50.0%; Pred. No. 53;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 11 DPEFMSETDL 20
 |:|:|:|:|:
 Db 1 DYPSLSATDI 10

RESULT 12
 AAB51604
 ID AAB51604 standard; peptide; 22 AA.
 XX
 AC AAB51604;
 XX
 DT 15-FEB-2001 (first entry)
 XX
 DE Yada homologous peptide #67.
 XX
 KW Proteobacteria; extracellular domain; virulence determinant; Yada;
 KW adhesin; proteobacterial infection prevention; vaccine.
 XX
 OS Actinobacillus actinomycetemcomitans.
 XX
 PN WO200061165-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 13-APR-2000; 2000WO-US009866.
 XX
 PR 13-APR-1999; 99US-0129073P.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PI Lupas AN;
 XX
 DR WPI; 2000-647397/62.
 XX
 PT An isolated polypeptide conserved in proteobacterial extracellular
 PT domains used in the treatment and prevention of bacterial infections.
 XX
 PS Example 5; Page 60; 85pp; English.

XX This invention relates to peptides AAB51512 - AAB51537 which represent
 CC conserved proteobacterial extracellular domains. Sequences AAB51538 -
 CC AAB51618 represent peptides homologous to Yada, a Yersinia adhesin which
 CC is an important virulence determinant of the Yersinia species. The
 CC invention includes an antibody which binds to the proteobacterial
 CC extracellular peptides, and an immunogenic composition containing the
 CC antibody used as a vaccine to prevent infection by a proteobacteria. The
 CC polypeptides and antibodies are useful in the treatment and prevention of
 CC proteobacterial infections. The polypeptides can also be used to identify
 CC compounds which antagonize the binding of a bacterial adhesin to its
 CC ligand. The host cell can be used to produce the polypeptides in a
 CC suitable culture system. The composition can be used to vaccinate a
 CC patient against a proteobacterial infection
 XX
 SQ Sequence 22 AA;

Query Match 34.7%; Score 35; DB 3; Length 22;
 Best Local Similarity 80.0%; Pred. No. 1.3e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 IAANSKDFPN 14

Db 9 IAANSKDAVN 18
 |||||

RESULT 13
 AAB51532
 ID AAB51532 standard; peptide; 24 AA.
 XX
 AC AAB51532;
 XX
 DT 15-FEB-2001 (first entry)
 XX

Proteobacterial extracellular conserved peptide SEQ ID 21.

XX Proteobacteria; extracellular domain; virulence determinant; Yada;
 KW adhesin; proteobacterial infection prevention; vaccine.
 XX
 OS Actinobacillus actinomycetemcomitans.
 XX
 PN WO200061165-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 13-APR-2000; 2000WO-US009866.
 XX
 PR 13-APR-1999; 99US-0129073P.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Lupas AN;
 XX
 DR WPI; 2000-647397/62.
 XX
 PT An isolated polypeptide conserved in proteobacterial extracellular
 PT domains used in the treatment and prevention of bacterial infections.
 XX
 PS Claim 8; Page 65; 85pp; English.

XX This invention relates to peptides AAB51512 - AAB51537 which represent
 CC conserved proteobacterial extracellular domains. Sequences AAB51538 -
 CC AAB51618 represent peptides homologous to Yada, a Yersinia adhesin which
 CC is an important virulence determinant of the Yersinia species. The
 CC invention includes an antibody which binds to the proteobacterial
 CC extracellular peptides, and an immunogenic composition containing the
 CC antibody used as a vaccine to prevent infection by a proteobacteria. The
 CC polypeptides and antibodies are useful in the treatment and prevention of
 CC proteobacterial infections. The polypeptides can also be used to identify
 CC compounds which antagonize the binding of a bacterial adhesin to its
 CC ligand. The host cell can be used to produce the polypeptides in a
 CC suitable culture system. The composition can be used to vaccinate a
 CC patient against a proteobacterial infection
 XX
 SQ Sequence 24 AA;

Query Match 34.7%; Score 35; DB 3; Length 24;
 Best Local Similarity 80.0%; Pred. No. 1.4e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 IAANSKDFPN 14
 |||||
 Db 11 IAANSKDAVN 20

RESULT 14
 AAB51532
 ID AAB51532 standard; protein; 29 AA.
 XX
 AC AAB51532;
 XX
 DT 25-MAR-2003 (revised)
 DT 06-MAR-1995 (first entry)
 XX

DE Ala-103 Lathyrus ochrus lectin (amino acids 87-115).
 XX legume lectin; anti-nutritional protein; surface loop; mutein;
 KW amino acid substitution; ligand-binding site.
 XX Lathyrus ochrus.

XX Key Location/Qualifiers
 FH Region 1..29
 FT Misc-difference 17 /note= "corresponds to amino acids 87-115 of pea lectin"
 FT /note= "wild-type Val replaced by Ala"

XX W09416088-A1.
 XX 21-JUL-1994.
 XX 14-JAN-1994; 94WO-NL000011.
 XX 15-JAN-1993; 93NL-00000078.

XX (UYLE-) RIJKSUNIV LEIDEN.
 XX (NEDE) NEDERLAND ORG TNO.
 XX Hoedemaeker PJ, Van Eijsden RR, Diaz ACL, De Pater B, Kijne JW;
 XX WPI; 1994-249234/30.

XX Improved legume lectin having reduced anti-nutritional effect - is prepd.
 PT by substituting residues in the ligand binding site of the protein, for
 PT use in fodder.
 XX Claim 8; Page 6; 36pp; English.

XX The anti-nutritional effect of legume lectins can be reduced by
 CC substituting at least one amino acid in a surface loop of the protein
 CC where the surface loop forms part of the ligand binding site. The
 CC substitution is pref. in the surface loop comprising amino acids 87-115
 CC of pea lectin or the corresp. residues of other lectins. The preferred
 CC substitution is alanine replacing the wild-type residue at position 103.
 CC The specification includes the wild-type sequences of surface loops from
 CC various legume lectins; sequences AAR57371-R57384 have been made by
 CC replacing with alanine, the wild-type residue corresp. to position 103 of
 CC the P.sativum sequence. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 29 AA;
 Query Match 34.7%; Score 35; DB 2; Length 29;
 Best Local Similarity 53.8%; Pred. No. 1.8e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 6 AANSKDFNMSST 18
 Db 17 AFNSKDYDKTSQT 29

RESULT 15
 ADA08124
 ID ADA08124 standard; peptide; 12 AA.
 AC ADA08124;

XX 06-NOV-2003 (first entry)

XX Alpha 1D adrenergic receptor peptide SEQ ID NO:26.
 XX identification; interacting protein; plurality; protein display library;
 KW binding protein; detection; specific binding;
 KW target-assisted iterative screening; protein-protein interaction.

OS Synthetic.
 OS Homo sapiens.

PN W02003029821-A1.

XX 10-APR-2003.

XX 01-OCT-2002; 2002WO-US031349.

XX 01-OCT-2001; 2001US-0326566P.

XX (BUCK-) BUCK INST.

XX Kourakine A, Bredesen D;

XX WPI; 2003-403031/38.

XX Identifying interacting proteins from a protein display library having
 PT potential binding proteins comprising contacting target proteins with the
 PT library and detecting binding of potential binding proteins with the
 PT target proteins.

XX Example 1; Page 29; 61pp; English.

XX The present invention describes a method for identifying interacting
 CC proteins from a plurality of potentially-interacting proteins comprising
 CC contacting one or more target proteins with a protein display library
 CC having a plurality of potential binding proteins and detecting specific
 CC binding of potential binding proteins with the target proteins. The
 CC method comprises: (a) contacting one or more target proteins with a
 CC protein display library comprising a plurality of potential binding
 CC proteins for one or more target proteins; (b) selecting members of the
 CC protein display library that bind to the target proteins to provide a
 CC preselected set of potential binding proteins; (c) separating the members
 CC of the preselected set of potential binding proteins from the bound
 CC target protein and immobilising the members on a solid support such that
 CC the members are spatially addressable; (d) contacting the members of the
 CC preselected set of potential binding proteins with the target proteins;
 CC and (e) detecting specific binding of members of the preselected set of
 CC potential binding proteins with the target proteins, where binding of a
 CC member of the preselected set with a target protein indicates that the
 CC member and the target protein are interacting proteins. Also described is
 CC a kit for carrying out the method described above, comprising a protein
 CC display library and instructional materials providing protocols for the
 CC above method. The target-assisted iterative screening method is useful
 CC for determining large numbers of interactions (e.g. protein-protein
 CC interactions) between members of a library and various targets. The
 CC present sequence represents a peptide, which is used in the
 CC exemplification of the present invention.

XX Sequence 12 AA;

Query Match 32.7%; Score 33; DB 6; Length 12;
 Best Local Similarity 50.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 11 DFPNMSETDL 20

Db 3 DYSNLRETDI 12

Search completed: July 15, 2004, 14:46:48
 Job time : 42.1333 secs

THIS PAGE BLANK (USPTO)

2000 JUL 1 10 51 AM '97

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:40:36 ; Search time 8.53333 Seconds
(without alignments)
225.449 Million cell updates/sec

Title: US-09-171-432A-45
Perfect score: 101
Sequence: 1 DLEETANNSKDPNMTDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 6282

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	32.7	24	2	female protein - h
2	33	32.7	26	2	beta-galactoside b
3	33	32.7	30	2	diuretic hormone I
4	30	29.7	29	2	hypothetical prote
5	29	28.7	19	2	phosphocarrier pro
6	29	28.7	20	2	phosphocarrier pro
7	28	27.7	28	2	interleukin-1 - st
8	27	26.7	19	2	15K protein A - ra
9	27	26.7	19	2	15K protein B - ra
10	26	25.7	19	2	capsid protein VP2
11	26	25.7	26	2	photosystem II D1
12	25	24.8	10	2	lg heavy chain C r
13	25	24.8	25	2	biphenyl dioxygena
14	25	24.8	27	2	GTP-binding regula
15	25	24.8	29	2	hypothetical prote
16	25	24.8	30	2	H4-exporting ATPas
17	25	24.8	30	2	H4-exporting ATPas
18	24	23.8	13	2	urinary tract ston
19	24	23.8	16	2	alpha-conotoxin Pn
20	24	23.8	23	2	glycine receptor a
21	24	23.8	26	2	arylacetyl-CoA N-a
22	24	23.8	28	2	apolipoprotein C-I
23	23	22.8	21	2	hypothetical prote
24	23	22.8	22	2	carboxypeptidase B
25	23	22.8	23	2	tyrosine kinase su
26	23	22.8	24	2	photosystem I 9.0K
27	23	22.8	26	2	tubulin alpha-1 ch
28	23	22.8	26	2	tubulin alpha-2 ch
29	23	22.8	27	2	superoxide dismuta

30 23 22.8 27 2 G97937 hypothetical prote
31 23 22.8 28 2 S08569 ribosomal protein
32 23 22.8 29 2 A48427 flavohemoglobin hm
33 23 22.8 30 2 S68639 nigroxin A - black
34 22 21.8 15 4 I38031 hypothetical MN1/T
35 22 21.8 16 2 PH0137 T-cell receptor be
36 22 21.8 20 2 S29817 cytochrome P450 2C
37 22 21.8 20 2 FN0115 insulin-like growth
38 22 21.8 20 2 A31049 caldesmon, fas
39 22 21.8 20 2 S57202 vitronectin-bindin
40 22 21.8 20 2 A48406 annexin VI homolog
41 22 21.8 20 2 C60822 cytochrome P450 UT
42 22 21.8 23 2 S37452 gene E6 protein (c
43 22 21.8 23 2 S37451 internal peptide V
44 22 21.8 23 2 B04348 ribosomal protein
45 22 21.8 24 2 B48401

ALIGNMENTS

RESULT 1

female protein - hamster (fragment)
C;Species: Cricetinae gen. sp. (hamster)
C;Date: 05-Jun-1987 #sequence_revision 02-Jun-1988 #text_change 30-Sep-1993
C;Accession: A19828
R;Coe, J.E.; Margossian, S.S.; Slayter, H.S.; Sogn, J.A.
J. Exp. Med. 153, 977-991, 1981

A;Title: Hamster female protein. A new pentraxin structurally and functionally similar to
A;Reference number: A19828; MUID:81241327; PMID:6166709
A;Accession: A19828
A;Molecule type: protein
A;Residues: 1-24 <COE>
C;Superfamily: C-reactive protein

Query Match 32.7%; Score 33; DB 2; Length 24;
Best Local Similarity 75.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 12 FPNMSETD 19
DB 9 FPROSETD 16

RESULT 2

beta-galactoside binding lectin - common marmoset (fragments)
C;Species: Callithrix jacchus (common marmoset)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C;Accession: PS0111
R;Ohsawa, F.; Hirano, F.; Natori, S.
J. Biochem. 107, 431-434, 1990

A;Title: Purification and properties of a beta-galactoside-binding lectin from neonatal
A;Reference number: PS0111; MUID:90256719; PMID:1692826
A;Accession: PS0111
A;Molecule type: protein
A;Residues: 1-26 <OHS>

Query Match 32.7%; Score 33; DB 2; Length 26;
Best Local Similarity 46.7%; Pred. No. 90;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 DLEETANNSKDPNMTDL 15
DB 6 DANTIVXNSKDSNNL 20

RESULT 3

diuretic hormone II - tobacco hornworm
N;Alternate names: diuretic peptide II
C;Species: Manduca sexta (tobacco hornworm)

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Apr-2001

C;Accession: J50645
R;Blackburn, M.B.; Kingan, T.G.; Bodnar, W.; Shabanowitz, J.; Hunt, D.F.; Kempe, T.; Wag
Biochem. Biophys. Res. Commun. 181, 927-932, 1991
A;Title: Isolation and identification of a new diuretic peptide from the tobacco hornworm
A;Reference number: J50645; MUID:92109782; PMID:1764106
C;Accession: J50645

A;Molecule type: protein
A;Residues: 1-30 <BLA>
A;Comment: This hormone is a factor which increases the rate of fluid excretion in the
C;Keywords: amidated carboxyl end
F;30/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 32.7%; Score 33; DB 2; Length 30;
Best Local Similarity 38.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 2 LERIAANSKDPFN 14

Db 16 MEKVAQNRRFLN 28

RESULT 4

T37120

hypothetical protein SCJ4.37 - Streptomyces coelicolor

C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T37120
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.

submitted to the EMBL Data Library, August 1999

A;Reference number: Z21588

A;Accession: T37120

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-29 <SAU>

A;Cross-references: EMBL:AL109950; PIDN: CAB52971.1; GSPDB: GN00070; SCODEB: SCJ4.37

A;Experimental source: strain A3(2)

C;Genetics: SCODEB: SCJ4.37

Query Match 29.7%; Score 30; DB 2; Length 29;

Best Local Similarity 43.8%; Pred. No. 3.1e+02;

Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 EIAANSKDPFNSETD 19

Db 4 ELLAKARLHPSETD 19

RESULT 5

A48400

phosphocarrier protein 1, Hpr-1 - Streptococcus salivarius (fragment)

C;Species: Streptococcus salivarius

C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 23-May-1997

C;Accession: A48400

R;Vadeboncoeur, C.; Konishi, Y.; Dumas, F.; Gauthier, L.; Frenette, M.

Biochimie 73, 1427-1430, 1991

A;Title: Hpr polymorphism in oral streptococci is caused by the partial removal of the N

A;Reference number: A48400; MUID:92190346; PMID:1799636

A;Accession: A48400

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-19 <VAD>

A;Note: sequence extracted from NCBI backbone (NCBIP:88289)

C;Superfamily: phosphotransferase system phosphohistidine-containing protein; phosphotra

Query Match 28.7%; Score 29; DB 2; Length 19;

Best Local Similarity 50.0%; Pred. No. 2.7e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 SKDPFNMSSETDL 20

Db 2 SKDPFHVAETGI 13

RESULT 6

B48400

phosphocarrier protein 2, Hpr-2 - Streptococcus salivarius (fragment)

C;Species: Streptococcus salivarius

C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 23-May-1997

C;Accession: B48400

R;Vadeboncoeur, C.; Konishi, Y.; Dumas, F.; Gauthier, L.; Frenette, M.

Biochimie 73, 1427-1430, 1991

A;Title: Hpr polymorphism in oral streptococci is caused by the partial removal of the N

A;Reference number: A48400; MUID:92190346; PMID:1799636

A;Accession: B48400

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-20 <VAD>

A;Note: sequence extracted from NCBI backbone (NCBIP:88290)

C;Superfamily: phosphotransferase system phosphohistidine-containing protein; phosphotra

C;Keywords: phosphoprotein

Query Match 28.7%; Score 29; DB 2; Length 20;

Best Local Similarity 50.0%; Pred. No. 2.9e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 SKDPFNMSSETDL 20

Db 3 SKDPFHVAETGI 14

RESULT 7

A61273

interleukin-1 - starfish (Asterias forbesii) (fragment)

C;Species: Asterias forbesii (Forbes's starfish)

C;Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 12-May-1994

C;Accession: A61273

R;Beck, G.; Habicht, G.S.

Mol. Immunol. 28, 577-584, 1991

A;Title: Purification and biochemical characterization of an invertebrate interleukin

A;Reference number: A61273; MUID:91319151; PMID:1861678

A;Accession: A61273

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-28 <BEC>

Query Match 27.7%; Score 28; DB 2; Length 28;

Best Local Similarity 40.0%; Pred. No. 6.2e+02;

Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 5 IAANSKDPFNMSSETD 19

Db 6 IANAKAYANAGRTD 20

RESULT 8

A38382

15K protein A - rabbit (fragment)

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 02-Sep-2000

C;Accession: A38382

R;Ooi, C.E.; Weiss, J.; Levy, O.; Elsbach, P.

J. Biol. Chem. 265, 15956-15962, 1990

A;Title: Isolation of two isoforms of a novel 15-kDa protein from rabbit polymorphonuc

A;Reference number: A38382; MUID:90368817; PMID:2203792

A;Accession: A38382

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-19 <OOI>

C;Superfamily: cathelin; cystatin homology

Query Match 26.7%; Score 27; DB 2; Length 19;

Best Local Similarity 41.7%; Pred. No. 5.7e+02;

Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

A:Accession: I54247
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-27 <RES>
A:Cross-references: GB:S70567; NID:G545549; PIDN:ADI4063.1; PID:G4261763
C:Superfamily: GTP-binding regulatory protein Gs alpha chain

Query Match 24.8%; Score 25; DB 2; Length 27;
Best Local Similarity 29.4%; Pred. No. 1.8e+03;
Matches 5; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 ERIAAKSDFPNMSETD 19
Db 10 KSLAKRSKELEKKLOED 26

RESULT 15
AC0717

Hypothetical protein STY1874b [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0717
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Mature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Simmonds, M.; Skelton, J.; Stevens, K.; Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0717
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-29 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02108.1; PID:G16502943; GSPDB:GN00176
C:Genetics:
A:Gene: STY1874b

Query Match 24.8%; Score 25; DB 2; Length 29;
Best Local Similarity 45.5%; Pred. No. 1.9e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 KDFPNMSETDL 20
Db 16 KAYYNLSITDI 26

Search completed: July 15, 2004, 14:52:13
Job time : 8.53333 secs

C;Keywords: chloroplast

Query Match 25.7%; Score 26; DB 2; Length 26;
Best Local Similarity 41.2%; Pred. No. 1.2e+03;
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 4 EIAANSKDFPNMGETDL 20
||| : ||
Db 10 EAAANILQDPDGXSVDL 26

RESULT 12

A37268
I9 heavy chain C region (129) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 19-Mar-1997 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998

C;Accession: A37268

R;Ruff-Jamison, S.; Campos-Gonzalez, R.; Glenney Jr., J.R.

J. Biol. Chem. 266, 6607-6613, 1991

A;Title: Heavy and light chain variable region sequences and antibody properties of anti

A;Reference number: A38740; MUID:91177923; PMID:1706720

A;Accession: A37268

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-10 <RUP>

Query Match 24.8%; Score 25; DB 2; Length 10;
Best Local Similarity 57.1%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 SKDFPNM 15
||: |||:
Db 2 SQSFPNV 8

RESULT 13

PN0632
biphenyl dioxygenase (EC 1.13.-.-) A4 - Pseudomonas sp. (strain LB400) (fragment)

C;Species: Pseudomonas sp.

C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 12-Apr-1995

C;Accession: PN0632

R;Hofer, B.; Ellis, L.D.; Dowling, D.N.; Timmis, K.N.

Gene 130, 47-55, 1993

A;Title: Genetic analysis of a Pseudomonas locus encoding a pathway for biphenyl/polychlor

A;Reference number: PN0632; MUID:93345822; PMID:8444527

A;Accession: PN0632

A;Molecule type: DNA

A;Residues: 1-25 <HOF>

A;Cross-references: GB:X66122

C;Genetics:

A;Gene: bphA4

C;Superfamily: toluene dioxygenase ferredoxin reductase component

C;Keywords: oxidoreductase; PCB biodegradation

Query Match 24.8%; Score 25; DB 2; Length 25;
Best Local Similarity 29.4%; Pred. No. 1.6e+03;
Matches 5; Conservative 5; Mismatches 3; Indels 4; Gaps 1;

QY 2 LEEIAANSKDFP---PN 14
||: ||:
Db 9 LQDVGSNLRDLKAKPN 25

RESULT 14

I54247

GTP-binding regulatory protein Gt alpha-2 chain - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000

C;Accession: I54247

R;Hirano, T.

Hokkaido Igaku Zasshi 68, 885-893, 1993

A;Title: [Analysis of cell specific transcription of the human cone transducin alpha su

A;Reference number: I54247; MUID:94156312; PMID:8112713

QY 3 EEIAANSKDFPN 14
||: ||:
Db 8 EEVVAQALQFEN 19

RESULT 9

B38382

15K protein B - rabbit (fragment)

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 02-Sep-2000

C;Accession: B38382

R;Odi, C.E.; Weiss, J.; Levy, O.; Elsbach, P.

J. Biol. Chem. 265, 15956-15962, 1990

A;Title: Isolation of two isoforms of a novel 15-kDa protein from rabbit polymorphonucle

A;Reference number: A38382; MUID:90368817; PMID:2203792

A;Accession: B38382

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-20 <OOI>

C;Superfamily: cathelin; cystatin homology

Query Match 26.7%; Score 27; DB 2; Length 20;
Best Local Similarity 41.7%; Pred. No. 6e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 EEIAANSKDFPN 14
||: ||:
Db 9 EEVVAQALQFEN 20

RESULT 10

PQ0548

capsid protein VP26 - human herpesvirus 1 (fragment)

C;Species: human herpesvirus 1

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999

C;Accession: PQ0548

R;Davidson, M.D.; Rixon, F.J.; Davison, A.J.

J. Gen. Virol. 73, 2709-2713, 1992

A;Title: Identification of genes encoding two capsid proteins (VP24 and VP26) of herpes

A;Reference number: PQ0544; MUID:93019027; PMID:1328483

A;Accession: PQ0548

A;Molecule type: protein

A;Residues: 1-19 <DAV>

A;Experimental source: strain 17

C;Genetics:

A;Gene: UL35

C;Keywords: capsid protein

Query Match 25.7%; Score 26; DB 2; Length 19;
Best Local Similarity 35.7%; Pred. No. 8.2e+02;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 5 IAANSKDFPNMSET 18
||: ||:
Db 5 LATNNQDNPHPQGT 18

RESULT 11

S65140

photosystem II D1 protein processing enzyme - spinach (fragments)

C;Species: Spinacia oleracea (spinach)

C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997

C;Accession: S65140

R;Inagaki, N.; Yamamoto, Y.; Mori, H.; Satoh, K.

Plant Mol. Biol. 30, 39-50, 1996

A;Title: Carboxyl-terminal processing protease for the D1 precursor protein: cloning and

A;Reference number: S65140; MUID:96197397; PMID:8616242

A;Accession: S65140

A;Status: preliminary

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:01 ; Search time 5.3333 Seconds
(without alignments)
195.263 Million cell updates/sec

Title: US-09-171-432A-45
Perfect score: 101
Sequence: 1 DLEETAAASKDFPNMSETDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues 2073
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	32.7	30	1 DIU2_HYLLI	P82015 hyles linea
2	33	32.7	30	1 DIU2_MANSE	P24858 manduca sex
3	25	24.8	30	1 FREA_LITIN	P82023 litoria inf
4	25	24.8	30	1 ACBI_DIGLA	P81624 digitalis l
5	25	24.8	30	1 BIF4_RABIT	P83470 oryctolagus
6	25	24.8	30	1 VAA1_EQUAR	Q04237 psilotum a
7	25	24.8	30	1 VAA1_PSINU	Q04237 psilotum nu
8	25	24.8	30	1 VAA2_PSINU	Q04239 psilotum nu
9	24	23.8	16	1 CXAA_CONPE	P50984 conus penna
10	24	23.8	28	1 APC1_RABIT	P33047 oryctolagus
11	23	22.8	20	1 HGL1_FASHE	P80527 fasciola he
12	23	22.8	21	1 YFLA_METVO	P42017 methanococc
13	23	22.8	27	1 RL5_HALHA	P50556 halobacteri
14	23	22.8	27	1 SODM_DESDE	P11419 desulfovibr
15	23	22.8	28	1 RL5_HALCU	P05972 halobacteri
16	23	22.8	30	1 TAT_HV1ZH	P12512 human immun
17	22	21.8	18	1 UC21_MAIZE	P80627 zea mays (m
18	22	21.8	19	1 TPIS_CLOPA	P81348 clostridium
19	22	21.8	20	1 CAQ1_RAT	P19633 rattus norv
20	22	21.8	21	1 LPRM_CORDI	P21232 corynebacte
21	22	21.8	22	1 23KD_BACST	P80166 bacillus st
22	22	21.8	23	1 VG22_BPT6	P1597 bacterioph
23	22	21.8	24	1 IRBP_SHEEP	P12663 ovis aries
24	22	21.8	25	1 GRP_SCYCA	P09472 scyllorhinu
25	22	21.8	28	1 LPFS_ECOLI	P22183 escherichia
26	22	21.8	29	1 RS7_METTE	O91639 methanosarc
27	22	21.8	30	1 MMAL_DESMI	P16312 dermatophag
28	21	20.8	10	1 UPA4_HUMAN	P30090 homo sapien
29	21	20.8	12	1 PA21_MICFM	P25072 micurus fu
30	21	20.8	12	1 V23K_WSSV	P82005 white spot
31	21	20.8	19	1 DHAB_COMTE	P80704 comamonas t
32	21	20.8	19	1 HBB2_UROHA	P18992 uromastyx h
33	21	20.8	20	1 F1BB_SHEEP	P14470 ovis aries

RESULT 1
DIU2_HYLLI
ID DIU2_HYLLI STANDARD; PRT; 30 AA.
AC P82015;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Diuretic hormone 2 (DH-2) (Diuretic peptide 2) (DP-2) (DH(30)).
OS Hyles lineata (Whitelined sphinx moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;
OC Sphingidae; Macroglossinae; Hyles.
OX NCBI_TaxID=103890;
RN [1]
RP SEQUENCE, ACTIVITY, AND AMIDATION.
RC TISSUE=Head;
RX MEDLINE=20161578; PubMed=10696588;
RA Furuya K., Harper M.A., Schegg K.M., Shooley D.A.;
RT "Isolation and characterization of CFR-related diuretic hormones from
the whitelined sphinx moth Hyles lineata."
RL Insect Biochem. Mol. Biol. 20:127-133(2000).
CC -!- FUNCTION: Regulation of fluid secretion. May stimulate primary
urine secretion by Malpighian tubules and causes a dose-dependent
stimulation of CAMP levels in the tubules.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=3574; METHOD=MALDI.
CC -!- SIMILARITY: Belongs to the sauvagine/corticotropin-releasing
factor/urotensin I family.
DR InterPro; IPR000187; corticoliberin.
DR Pfam; PF00473; CRF; 1.
DR SMART; SM00039; CRF; 1.
DR PROSITE; PS00511; CRF; FALSE_NEG.
KW Hormone; Amidation.
FT MOD_RES 30
SQ SEQUENCE 30 AA; 3575 MW; 5C2D6BD2DD8BFC67 CRC64;
Query Match 32.7%; Score 33; DB 1; Length 30;
Best Local Similarity 38.5%; Pred. No. 89;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LEEITAANSKDFPN 14
Db 16 MEKVAQNRRNFLN 28
RESULT 2
DIU2_MANSE
ID DIU2_MANSE STANDARD; PRT; 30 AA.
AC P24858;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Diuretic hormone 2 (DH-2) (Diuretic peptide 2) (DP-2) (DPI1).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;

DR InterPro: IPR001089; CXC chmkine smil.
 DR PROSITE; PS00471; SMALL_CYTOKINES_CXC; PARTIAL.
 KW Cytokine; Heparin-binding; Chemotaxis.
 FT VARIANT 20 20 R -> S.
 FT VARIANT 25 25 G -> T.
 FT NON TER 30 30
 SQ SEQUENCE 30 AA; 3317 MW; 70C6D72CDA20DF98 CRC64;

Query Match 24.8%; Score 25; DB 1; Length 30;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 11 DFFNMSETDL 20
 Db 2 DDPKSEGGDL 11

RESULT 6
 VAA1 EQUAR STANDARD; PRT; 30 AA.
 ID VAA1 EQUAR
 AC Q04236;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vacuolar ATP synthase catalytic subunit A, isoform 1 (EC 3.6.3.14) (Fragment).
 DE (Fragment).
 OS Equisetum arvense (Field horsetail) (Common horsetail).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Moniliformopses; Equisetophyta; Sphenopsida; Equisetales;
 OC Equisetaceae; Equisetum.
 OX NCBI_TaxID=3258;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93138084; PubMed=8422915;
 RA Starke T., Gogarten J.P.;
 RT "A conserved intron in the V-ATPase A subunit genes of plants and algae."
 RL FEBS Lett. 315:252-258(1993).
 CC -!- FUNCTION: Catalytic subunit of the peripheral V1 complex of vacuolar ATPase. V-ATPase vacuolar ATPase is responsible for acidifying a variety of intracellular compartments in eukaryotic cells.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
 CC -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a peripheral catalytic V1 complex (main components: subunits A, B, C, D, E, and F) attached to an integral membrane V0 proton pore complex (main component: the proteolipid protein).
 CC -!- MISCELLANEOUS: TWO SEPARATE GENES ENCODE THE CATALYTIC 70 kDa V-ATPASE SUBUNIT IN PSILOTUM AND Equisetum.
 CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; X56983; CAA40301.1; -
 CC PIR; S21814; S21814.
 DR InterPro; IPR000194; ATPase_a/bcentre.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; PARTIAL.
 KW ATP synthesis; Hydrogen ion transport; Hydrolase; ATP-binding; Multigene family.
 FT NON TER 1 1
 FT NON TER 30 30
 SQ SEQUENCE 30 AA; 3337 MW; 9627ED62068D761F CRC64;

Query Match 24.8%; Score 25; DB 1; Length 30;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 11 DFFNMSETDL 20
 Db 2 DDPKSEGGDL 11

RESULT 8
 VAA2 PSINU STANDARD; PRT; 30 AA.
 ID VAA2 PSINU
 AC Q04239;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

Qy 11 DFFNMSET 18
 Db 6 DFFQLTWT 13

RESULT 7
 VAA1 PSINU STANDARD; PRT; 30 AA.
 ID VAA1 PSINU
 AC Q04237;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vacuolar ATP synthase catalytic subunit A, isoform 1 (EC 3.6.3.14) (Fragment).
 DE (Fragment).
 OS Psilotum nudum (Whisk fern).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Moniliformopses; Psilotophyta; Psilotales; Psilotaceae; Psilotum.
 OX NCBI_TaxID=3240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93138084; PubMed=8422915;
 RA Starke T., Gogarten J.P.;
 RT "A conserved intron in the V-ATPase A subunit genes of plants and algae."
 RL FEBS Lett. 315:252-258(1993).
 CC -!- FUNCTION: Catalytic subunit of the peripheral V1 complex of vacuolar ATPase. V-ATPase vacuolar ATPase is responsible for acidifying a variety of intracellular compartments in eukaryotic cells.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate + H(+) (Out).
 CC -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a peripheral catalytic V1 complex (main components: subunits A, B, C, D, E, and F) attached to an integral membrane V0 proton pore complex (main component: the proteolipid protein).
 CC -!- MISCELLANEOUS: TWO SEPARATE GENES ENCODE THE CATALYTIC 70 kDa V-ATPASE SUBUNIT IN PSILOTUM AND Equisetum.
 CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; X56985; CAA40303.1; -
 CC PIR; S21816; S21816.
 DR InterPro; IPR000194; ATPase_a/bcentre.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; PARTIAL.
 KW ATP synthesis; Hydrogen ion transport; Hydrolase; ATP-binding; Multigene family.
 FT NON TER 1 1
 FT NON TER 30 30
 SQ SEQUENCE 30 AA; 3380 MW; 9627ED62069E561F CRC64;

Query Match 24.8%; Score 25; DB 1; Length 30;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 11 DFFNMSET 18
 Db 6 DFFQLTWT 13

RESULT 8
 VAA2 PSINU STANDARD; PRT; 30 AA.
 ID VAA2 PSINU
 AC Q04239;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vacuolar ATP synthase catalytic subunit A, isoform 2 (EC 3.6.3.14)
 DE (Fragment).
 OS Psilotum nudum (Whisk fern).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Moniliformopses; Psilotophyta; Psilotales; Psilotaceae; Psilotum.
 OX NCBI_TaxID=3240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93138084; PubMed=8422915;
 RA Starke T., Gogarten J.P.;
 RT "A conserved intron in the V-ATPase A subunit genes of plants and
 RT algae";
 RL FEBS Lett. 315:252-258(1993).
 CC -!- FUNCTION: Catalytic subunit of the peripheral V1 complex of a
 CC vacuolar ATPase. V-ATPase vacuolar ATPase is responsible for
 CC acidifying a variety of intracellular compartments in eukaryotic
 CC cells.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -!- SUBUNIT: V-ATPase is an heteromultimeric enzyme composed of a
 CC peripheral catalytic V1 complex (main components: subunits A, B,
 CC C, D, E, and F) attached to an integral membrane V0 proton pore
 CC complex (main component: the proteolipid protein).
 CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; X56986; CAA40304.1; -.
 DR PIR; S21816; S21816.
 DR InterPro; IPR000194; ATPase.a/bcentre.
 DR PROSITE; PS00152; ATPase ALPHA.BETA; PARTIAL.
 KW ATP synthesis; Hydrogen ion transport; Hydrolase; ATP-binding;
 KW Multigene family.
 FT NON_TER 1 1
 FT TER 30 30
 SQ SEQUENCE 30 AA; 3380 MW; 9627ED62069E561F CRC64;

Query Match 24.8%; Score 25; DB 1; Length 30;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 DFPNMSET 18
 |||||
 Db 6 DFFQLTWT 13

RESULT 9
 CXXA CONPE STANDARD; PRT; 16 AA.
 ID CXXA CONPE STANDARD; PRT; 16 AA.
 AC P50984;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alpha-conotoxin PnIA.
 OS Conus pennaceus (Feathered cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=37335;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=94347719; PubMed=8068627;
 RA Fainzilber M., Hasson A., Oren R., Burlingame A.L., Gordon D.,
 RA Spira M.E., Zlotkin E.;
 RT "New mollusc-specific alpha-conotoxins block Aplysia neuronal

RT acetylcholine receptors.";
 RL Biochemistry 33:9523-9529(1994).
 RN [2]
 RP SULFATION OF TYR-15.
 RX MEDLINE=99242956; PubMed=10226369;
 RA Wolfender J.L., Chu F., Ball H., Wolfender F., Fainzilber M.,
 RA Baldwin M.A., Burlingame A.L.;
 RT "Identification of tyrosine sulfation in Conus pennaceus conotoxins
 RT alpha-PnIA and alpha-PnIB: further investigation of labile sulfo- and
 RT phosphopeptides by electrospray, matrix-assisted laser
 RT desorption/ionization (MALDI) and atmospheric pressure MALDI mass
 RT spectrometry.";
 RL J. Mass Spectrom. 34:447-454(1999).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS).
 RX MEDLINE=96311277; PubMed=8740364;
 RA Hu S.-H., Gehrmann J., Guddat L.W., Allewood P.F., Craik D.J.,
 RA Martin J.L.;
 RT "The 1.1 A crystal structure of the neuronal acetylcholine receptor
 RT antagonist, alpha-conotoxin PnIA from Conus pennaceus.";
 RL Structure 4:417-423(1996).
 CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
 CC bind to the nicotinic acetylcholine receptors (nAChR) and thus
 CC inhibit them. In contrast to other alpha-conotoxins, which are
 CC selective for vertebrate skeletal muscle nAChR, the Conus
 CC pennaceus alpha-conotoxins block nAChR in mollusks.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- SIMILARITY: Belongs to the conotoxin A-superfamily. Alpha-type
 CC family.
 DR PIR; A54877; A54877.
 DR PDB; 1PN; 21-APR-97.
 DR Postsynaptic neurotoxin; Neurotoxin; Toxin;
 KW Acetylcholine receptor inhibitor; Amidation; Sulfation; 3D-structure.
 FT DISULFID 2 8
 FT DISULFID 3 16
 FT MOD_RES 15 15 SULFATION.
 FT MOD_RES 16 16 AMIDATION.
 FT HELIX 2 4
 FT HELIX 6 11
 FT TURN 13 16
 SQ SEQUENCE 16 AA; 1628 MW; 05310FF95EC99005 CRC64;

Query Match 23.8%; Score 24; DB 1; Length 16;
 Best Local Similarity 45.5%; Pred. No. 1.2e+03;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 LEEIAANSKDF 12
 |||||
 Db 5 LPPCAANNPDY 15

RESULT 10
 APC1 RABIT STANDARD; PRT; 28 AA.
 ID APC1 RABIT STANDARD; PRT; 28 AA.
 AC P33047;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Apolipoprotein C-I (Apo-CI) (Fragment).
 GN APOC1.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91093092; PubMed=2266137;
 RA Weisgraber K.H., Mahley R.W., Kowal R.C., Herz J., Goldstein J.L.,
 RA Brown M.S.;
 RT "Apolipoprotein C-I modulates the interaction of apolipoprotein E
 RT with beta-migrating very low density lipoproteins (beta-VLDL) and
 RT inhibits binding of beta-VLDL to low density lipoprotein

```

RT receptor-related protein.";
RL J. Biol. Chem. 265:22453-22459(1990).
CC -!- FUNCTION: Appears to modulate the interaction of APOE with beta-
CC migrating VLDL and inhibit binding of beta-VLDL to the LDL
CC receptor-related protein.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Secreted in plasma.
CC -!- SIMILARITY: Belongs to the apolipoprotein C1 family.
DR PIR; A23691; A23691.
DR HSP; P02654; IALE.
DR InterPro; IPR006781; Apoc-I.
DR Pfam; PF04691; Apoc-I; 1.
KW Transport; Lipid transport; VLDL.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3165 MW; 2903C027F676541C CRC64;

Query Match 23.8%; Score 24; DB 1; Length 28;
Best Local Similarity 43.8%; Pred. No. 2e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LEEIAANSKDFPNMSE 17
||| ||| ||| |||
Db 7 LELIPDKLKEFGNTLE 22

RESULT 11
HGLI_FASHE STANDARD; PRT; 20 AA.
AC P80527;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemoglobinase-like protein 1 (EC 3.4.22.34) (Newly excysted juvenile
DE protein 3) (Fragment).
OS Fasciola hepatica (liver fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Echinostomida; Echinostomata; Fasciolidae; Fasciola.
OX NCBI_TaxID=6192;
[1]
RN SEQUENCE.
RX MEDLINE=95366993; PubMed=7639732;
RA Tkalcovic J., Ashman K., Meeusen E.;
RT "Fasciola hepatica: rapid identification of newly excysted juvenile
RT proteins.";
RL Biochem. Biophys. Res. Commun. 213:169-174(1995).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins and small-molecule
CC substrates at -Asn-|-Xaa- bonds.
CC -!- DEVELOPMENTAL STAGE: Expressed at the newly excysted juvenile
CC stage.
CC -!- SIMILARITY: Belongs to peptidase family C13.
DR MEROPS; C13.003; -.
DR InterPro; IPR001096; Peptidase C13.
DR Pfam; PF01650; Peptidase C13; 1.
DR KEGG; K01650; Thiol protease.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2278 MW; 865F368A1A03874D CRC64;

Query Match 22.8%; Score 23; DB 1; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 5 IAAANSKDFPN 14
: ||| : |||
Db 6 IVAGSNGWPN 15

RESULT 12
YFLA_METVO STANDARD; PRT; 21 AA.
AC P42017;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Hypothetical protein in flaB3 3'region (Fragment).
OS Methanococcus voltae.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=2188;
RN SEQUENCE FROM N.A.
RP STRAIN=DSM 1537 / PS;
RX MEDLINE=92041608; PubMed=1718944;
RA Kalmokoff M.L., Jarrell K.F.;
RT "Cloning and sequencing of a multigene family encoding the flagellins
RT of Methanococcus voltae.";
RL J. Bacteriol. 173:7113-7125(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M72148; -; NOT_ANNOTATED_CDS.
DR PIR; E41316; E41316.
KW Hypothetical protein.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2404 MW; 1FA9CC772369C2DC CRC64;

Query Match 22.8%; Score 23; DB 1; Length 21;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 13 PNMSE 17
||| |||
Db 2 PNISE 6

RESULT 13
RL5_HALHA STANDARD; PRT; 23 AA.
AC P50556;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S ribosomal protein L5P (HHA5) (Fragment).
GN RPL5P.
OS Halobacterium halobium.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RN SEQUENCE.
RP STRAIN=DSM 670;
RX MEDLINE=94229075; PubMed=8174557;
RA McDougall J., Wittmann-Liebold B.;
RT "Comparative analysis of the protein components from 5S rRNA, protein
RT complexes of halophilic archaeobacteria.";
RL Eur. J. Biochem. 221:779-785(1994).
CC -!- SIMILARITY: Belongs to the L5P family of ribosomal proteins.
DR InterPro; IPR002132; Ribosomal L5.
DR Pfam; PF00281; Ribosomal L5; 1.
DR PROSITE; PS00358; RIBOSOMAL_L5; PARTIAL.
KW Ribosomal protein.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2773 MW; 19B318335A19B298 CRC64;

Query Match 22.8%; Score 23; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 8 NSKDFPNMSE 17
: ||| : |||
Db 4 DSTDFHEMRE 13

```

```
RESULT 14
SODM DESDE
ID SODM_DESDE STANDARD; PRT; 27 AA.
AC FL1419;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Superoxide dismutase [Mn] (EC 1.15.1.1) (Fragments).
GN SODA.
OS Desulfovibrio desulfuricans.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=876;
RN (1)
RP SEQUENCE.
RX MEDLINE=77162619; PubMed=323056;
RA Bruschi M., Hatchikian C., Bonicel J., Bovier-Lapierre G.,
RA Couchoud P.;
RT "The N-terminal sequence of superoxide dismutase from the strict
RT anaerobe Desulfovibrio desulfuricans.";
RL FEBS Lett. 76:121-124(1977).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 manganese ion per subunit (By similarity).
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: Belongs to the iron/manganese superoxide dismutase
CC family.
DR PIR; A12968; A12968.
DR InterPro; IPR001189; SODismutase.
DR PROSITE; PS00088; SOD MN; PARTIAL.
KW Oxidoreductase; Metal-binding; Manganese.
FT NON_CONS 15 16
FT NON_CONS 25 26
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2982 MW; 96D198209F672634 CRC64;

Query Match 22.8%; Score 23; DB 1; Length 27;
Best Local Similarity 45.5%; Pred. No. 2.8e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LEEIAANSKDF 12
Db 15 LPKISAKTFDF 25

RESULT 15
RL5 HALCU
ID RL5 HALCU STANDARD; PRT; 28 AA.
AC P05972;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE S0S ribosomal protein L5P (HCU15) (Fragment).
GN RPL5P.
OS Halobacterium cutirubrum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN (1)
RP SEQUENCE.
RX MEDLINE=79045279; PubMed=152199;
RA Smith N., Matheson A.T., Yaguchi M., Willick G., Nazar R.N.;
RT "The 5-S RNA-protein complex from an extreme halophile,
RT Halobacterium cutirubrum. Purification and characterization.";
RL Eur. J. Biochem. 89:501-509(1978).
CC -!- SIMILARITY: Belongs to the L5P family of ribosomal proteins.
DR InterPro; IPR002132; Ribosomal L5.
DR Pfam; PF00281; Ribosomal L5; 1
DR PROSITE; PS00358; RIBOSOMAL_L5; PARTIAL.
```

```
KW Ribosomal protein.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3172 MW; 47CF2FA46689B318 CRC64;

Query Match 22.8%; Score 23; DB 1; Length 28;
Best Local Similarity 50.0%; Pred. No. 2.9e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 NSKDFENMSE 17
Db 4 DSTDFHEMRE 13
```

Search completed: July 15, 2004, 14:47:34
Job time : 6.33333 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	35	34.7	29	11	Q8BP29	Q8bp29 mus musculus
2	34	33.7	15	6	Q3TQ09	Q9tqd9 bos taurus
3	33	32.7	30	15	P90307	P90307 human immun
4	31	30.7	28	13	Q5Z2F5	Q5z2f5 gallus gall
5	30	29.7	24	2	O52389	O52389 anabaena fl
6	30	29.7	29	5	Q9GU41	Q9gu41 scypha raph
7	30	29.7	29	16	Q9S1T6	Q9s1t6 streptomyce
8	30	29.7	30	16	Q7VNX3	Q7vnx3 haemophilus
9	29	28.7	15	6	Q9TR40	Q9tr40 bos taurus
10	29	28.7	22	5	Q95V15	Q95vi5 heliconius
11	29	28.7	22	5	Q95NX1	Q95nx1 heliconius
12	29	28.7	22	5	Q95VJ2	Q95vj2 heliconius
13	29	28.7	22	5	Q95NX2	Q95nx2 heliconius
14	29	28.7	26	6	Q95J99	Q95j99 sus scrofa
15	29	28.7	28	2	Q9R514	Q9r5i4 streptococc
16	28	27.7	20	2	O51950	O51950 chlmydia t

AC Q9TC09;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glutamate dehydrogenase isoform I (EC 1.4.1.2) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RN [2]
RP SEQUENCE
RX MEDLINE=96061967; PubMed=7588764;
RA Cho S.W., Lee J., Choi S.Y.;
RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine brain.";
RL Eur. J. Biochem. 233:340-346 (1995).
RN [2]
RP SEQUENCE
RX MEDLINE=96043916; PubMed=7581004;
RA Lee J., Kim S.W., Cho S.W.;
RT "A novel glutamate dehydrogenase from bovine brain: purification and characterization.";
RL Biochem. Mol. Biol. Int. 36:1087-1096 (1995).
DR GO:0004352; F:glutamate dehydrogenase activity; IEA.
DR GO:0004352; F:glutamate dehydrogenase activity; IEA.
SQ SEQUENCE 15 AA; 1754 MW; 65F7CD31023AEEBA CRC64;

Query Match 33.7%; Score 34; DB 6; Length 15;
Best Local Similarity 58.3%; Pred. No. 2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 BETAANSKDPFN 14
DB 1 EEAADREDDPN 12

RESULT 3
P90307 ID P90307 PRELIMINARY; PRT; 30 AA.
AC P90307;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Surface glycoprotein gp120 (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96022623; PubMed=9359665;
RA Roth W.W., Levett P.N., Hudson C.P., Roach T.C., Womack C., Bond V.C.;
RT "HIV type 1 envelope sequences from seroconverting patients in Barbados.";
RL AIDS Res. Hum. Retroviruses 13:1443-1446 (1997).
DR EMBL: U80240; AB38694.1; -;
DR GO:0019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005198; F:structural molecule activity; IEA.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein.
FT NON TER 1
SQ SEQUENCE 30 AA; 3272 MW; B6E93BC48FCA0ED5 CRC64;

Query Match 32.7%; Score 33; DB 15; Length 30;
Best Local Similarity 37.5%; Pred. No. 5.8e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 BETAANSKDPFNSET 18
DB 9 BEVIRSKFTDNKNT 24

RESULT 4
Q7SZF5 ID Q7SZF5 PRELIMINARY; PRT; 28 AA.
AC Q7SZF5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aggrecan chondroitin sulfate CS2 domain (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=nanomelia;
RX MEDLINE=95128519; PubMed=7827752;
RA Primorac D., Stover M.L., Clark S.H., Rowe D.W.;
RT "Molecular basis of nanomelia, a heritable chondrodystrophy of chicken.";
RL Matrix Biol. 14:297-305 (1994).
DR EMBL: S74657; AAC60753.2; -;
DR NON TER 1
DR NON TER 28
FT NON TER 28
SQ SEQUENCE 28 AA; 2986 MW; 802E5A9ED6A7A098 CRC64;

Query Match 30.7%; Score 31; DB 13; Length 28;
Best Local Similarity 35.7%; Pred. No. 1.1e+03;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 BETAANSKDPFNMS 16
DB 6 QEVSGESSAPPEIS 19

RESULT 5
O52389 ID O52389 PRELIMINARY; PRT; 24 AA.
AC O52389;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine/threonine phosphatase 2A (EC 3.1.3.16) (Serine/threonine protein phosphatase) (Fragment).
GN P22A-CYANO2.
OS Anabaena flos-aquae.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=1166;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NRC-525-17;
RX MEDLINE=98047369; PubMed=9385145;
RA Shi L., Carmichael W.W.;
RT "pp1-cyano2, a protein serine/threonine phosphatase 1 gene from the cyanobacterium Microcystis aeruginosa UTEX 2063.";
RL Arch. Microbiol. 168:528-531 (1997).
CC -1- CATALYTIC ACTIVITY: A PHOSPHOPROTEIN + H(2)O = A PROTEIN + -1- PHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES.
DR EMBL: AF037160; AAC46048.1; -;
DR HSSP: P08129; 1FJM.
DR GO:0004723; F:calcium-dependent protein serine/threonine . . .; IEA.
DR GO:0008420; F:CTD phosphatase activity; IEA.
DR GO:0016787; F:hydrolase activity; IEA.
DR GO:0004724; F:magnesium-dependent protein serine/threonine . . .; IEA.
DR GO:0030145; F:manganese ion binding; IEA.
DR GO:0017018; F:myosin phosphatase activity; IEA.
DR GO:0000158; F:protein phosphatase type 2A activity; IEA.
DR GO:0030357; F:protein phosphatase type 2B activity; IEA.
DR GO:0015071; F:protein phosphatase type 2C activity; IEA.
DR InterPro: IPR006186; T_phase_apah.
DR Problem: PD000252; T_phase_apah; 1.


```

KW Hydrolase; Iron; Manganese.
FT NON TER 1
FT NON TER 24 24
SQ SEQUENCE 24 AA; 2812 MW; 774EA261FC0626C7 CRC64;

Query Match 29.7%; Score 30; DB 2; Length 24;
Best Local Similarity 46.2%; Pred. No. 1.3e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFP 13
DB 4 DLKELFKNGGDCP 16

RESULT 6
OSGU41 PRELIMINARY; PRT; 29 AA.
AC Q9GU41;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Homeobox protein NKX2 (Fragment).
GN NKX2.
OS Scypha raphanus.
OC Eukaryota; Metazoa; Porifera; Calcarea; Calcaronea; Leucosoleniida;
OC Sycettidae; Scypha.
OX NCBI_TaxID=56443;
RN [1]_TaxID=56443;
RP SEQUENCE FROM N.A.
RX MEDLINE=20476455; PubMed=11020308;
RA Manuel M., Le Parco Y.;
RT "Homeobox Gene Diversification in the Calcareous Sponge, Sycon raphanus."
RL Mol. Phylogenet. Evol. 17:97-107(2000).
DR EMBL; AF197114; AAG28514.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
FT NON TER 1
FT NON TER 29
SQ SEQUENCE 29 AA; 3443 MW; C359EDBD4D0FC332 CRC64;

Query Match 29.7%; Score 30; DB 5; Length 29;
Best Local Similarity 40.0%; Pred. No. 1.6e+03;
Matches 8; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 1 DLEETAAANSKDFP 20
DB 13 DREELARLK---LTETQI 28

RESULT 7
OSGU41 PRELIMINARY; PRT; 29 AA.
AC Q9S1T6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein SCO0056.
GN SCO0056 OR SGJ4.37.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]_
RP SEQUENCE FROM N.A.
RX STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

KW Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939104; CAB52971.1; -.
DR PIR; T37120; T37120.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 29 AA; 3254 MW; C97B50CAB5DA1D11 CRC64;

Query Match 29.7%; Score 30; DB 16; Length 29;
Best Local Similarity 43.8%; Pred. No. 1.6e+03;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 4 ETAAANSKDFP 19
DB 4 ELLAKARLHP 19

RESULT 8
OSGU41 PRELIMINARY; PRT; 30 AA.
AC Q7VNX3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN HD0352.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]_TaxID=730;
RP SEQUENCE FROM N.A.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S., Ji., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017151; AAP95324.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 30 AA; 3771 MW; FBC107200A668604 CRC64;

Query Match 29.7%; Score 30; DB 16; Length 30;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 8; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

QY 8 NSK-----DPPNMSE 17
DB 15 NSKKLEFLEPNYSE 30

RESULT 9
OSGU41 PRELIMINARY; PRT; 15 AA.
AC Q9TR40;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Glutamate dehydrogenase isoform II (EC 1.4.1.2) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]_
RP SEQUENCE.
RX MEDLINE=96061967; PubMed=7588764;
RA Cho S.W., Lee J., Choi S.Y.;

```

RT "Two soluble forms of glutamate dehydrogenase isoproteins from bovine
 brain";
 RL Eur. J. Biochem. 233:340-346(1995).
 DR GO; GO:0004352; F:glutamate dehydrogenase activity; IEA.
 SQ SEQUENCE 15 AA; 1724 MW; 65F7CD91023AF925 CRC64;

Query Match 28.7%; Score 29; DB 6; Length 15;
 Best Local Similarity 54.3%; Pred. NO. 1.2e+03;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 EIAANSKDPEN 14
 | | | | |
 Db 2 EAAADREDDEN 12

RESULT 10

Q95V15 ID Q95V15 PRELIMINARY; PRT; 22 AA.
 AC Q95V15; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Mannose phosphate isomerase (Fragment).
 GN MPI.
 OS Heliconius cydno chioneus.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Papilionoidea; Nymphalidae; Heliconiinae; Heliconius.
 OX NCBI_TaxID=171915;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=STRI-B-570-Mpi-1;
 RA Beltran M.S., Jiggins C.D., Bull V., McMillan O., Bermingham E.,
 RA Mallet J., Linhares M.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF413739; AAL07927.1; -.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 KW Isomerase.
 FT NON TER 1
 FT NON TER 22
 SQ SEQUENCE 22 AA; 2530 MW; 2DE4FB8E2C362CE6 CRC64;

Query Match 28.7%; Score 29; DB 5; Length 22;
 Best Local Similarity 83.3%; Pred. NO. 1.8e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 KDFPNM 15
 | | | | |
 Db 17 KSFPNM 22

RESULT 11

Q95NX1 ID Q95NX1 PRELIMINARY; PRT; 22 AA.
 AC Q95NX1; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Mannose phosphate isomerase (Fragment).
 GN MPI.
 OS Heliconius melpomene rosina.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Papilionoidea; Nymphalidae; Heliconiinae; Heliconius.
 OX NCBI_TaxID=171916;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=STRI-B-544-Mpi-1, and STRI-B-811-Mpi-1;
 RA Beltran M.S., Jiggins C.D., Bull V., McMillan O., Bermingham E.,
 RA Mallet J., Linhares M.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF413740; AAL07928.1; -.
 DR EMBL; AF413743; AAL07931.1; -.

DR GO; GO:0016853; F:isomerase activity; IEA.

KW Isomerase.
 FT NON TER 1
 FT NON TER 22
 SQ SEQUENCE 22 AA; 2530 MW; 2DE4FB8E2C362CE6 CRC64;

Query Match 28.7%; Score 29; DB 5; Length 22;
 Best Local Similarity 83.3%; Pred. NO. 1.8e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 KDFPNM 15
 | | | | |
 Db 17 KSFPNM 22

RESULT 12

Q95VJ2 ID Q95VJ2 PRELIMINARY; PRT; 22 AA.
 AC Q95VJ2; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Mannose phosphate isomerase (Fragment).
 GN MPI.
 OS Heliconius heurippa.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Papilionoidea; Nymphalidae; Heliconiinae; Heliconius.
 OX NCBI_TaxID=33423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=STRI-B-8-Mpi-2;
 RA Beltran M.S., Jiggins C.D., Bull V., McMillan O., Bermingham E.,
 RA Mallet J., Linhares M.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF413730; AAL07918.1; -.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 KW Isomerase.
 FT NON TER 1
 FT NON TER 22
 SQ SEQUENCE 22 AA; 2530 MW; 2DE4FB8E2C362CE6 CRC64;

Query Match 28.7%; Score 29; DB 5; Length 22;
 Best Local Similarity 83.3%; Pred. NO. 1.8e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 KDFPNM 15
 | | | | |
 Db 17 KSFPNM 22

RESULT 13

Q95NX2 ID Q95NX2 PRELIMINARY; PRT; 22 AA.
 AC Q95NX2; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Mannose phosphate isomerase (Fragment).
 GN MPI.
 OS Heliconius melpomene cythera.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Papilionoidea; Nymphalidae; Heliconiinae; Heliconius.
 OX NCBI_TaxID=171918;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=STRI-B-8074-Mpi-1, and STRI-B-8074-Mpi-2;
 RA Beltran M.S., Jiggins C.D., Bull V., McMillan O., Bermingham E.,
 RA Mallet J., Linhares M.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF413735; AAL07923.1; -.
 DR EMBL; AF413736; AAL07924.1; -.

DR GO:0016853; F:isomerase activity; IEA.
 KW Isomerase.
 FT NON_TER 1 1
 FT NON_TER 22 22
 SQ SEQUENCE 22 AA; 2530 MW; 2DE4FB8E2C362CE6 CRC64;
 Query Match 28.7%; Score 29; DB 5; Length 22;
 Best Local Similarity 83.3%; Pred. No. 1.8e+03;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 10 KDFPNM 15
 DB 17 KSFPNM 22
 RESULT 14
 Q95J99 PRELIMINARY; PRT; 26 AA.
 AC Q95J99;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Laminin beta 2 chain (Fragment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]_TaxID=9823;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cardiovascular;
 RA Markmann A., Kresse H.;
 RT "Regulation of VSMC Differentiation."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF330203; AAL09469.1; -.
 FT NON_TER 1 1
 FT NON_TER 26 26
 SQ SEQUENCE 26 AA; 2972 MW; 82E4C3F1F18C109 CRC64;
 Query Match 28.7%; Score 29; DB 6; Length 26;
 Best Local Similarity 31.6%; Pred. No. 2.1e+03;
 Matches 6; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
 QY 1 DLEIIAANSKDFPNMSETD 19
 :||: :||| :
 DB 7 ELRELISVKDFLSQEGAD 25
 RESULT 15
 Q9RS14 PRELIMINARY; PRT; 28 AA.
 AC Q9RS14;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE LOW-molecular weight protein, LMP-12KDA-4.7PI=PHOSPHOCARRIER protein homolog (Fragment).
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92393218; PubMed=1520957;
 RA Gerlach D., Alouf H., Moravsek L., Pavlik M., Kohler W.;
 RT "The characterization of two new low molecular weight proteins (LMPs) from Streptococcus pyogenes."
 RL Int. J. Med. Microbiol. Virol. Parasitol. Infect. Dis. 277:1-9(1992).
 DR HSSP; P07515; 1PTF.
 DR GO:0005351; F:sugar porter activity; IEA.
 DR GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
 DR GO:0006810; P:transport; IEA.
 DR InterPro; IPR001020; HPr HisP S.
 DR InterPro; IPR000032; HPr protein.

DR Pfam: PF00381; PTS-HPr; 1.
 DR ProDom; PD002338; HPr_protein; 1.
 DR PROSITE; PS00369; PTS_HPR_HIS; 1.
 FT NON_TER 1 1
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 2993 MW; 8892A9CB39941870 CRC64;
 Query Match 28.7%; Score 29; DB 2; Length 28;
 Best Local Similarity 50.0%; Pred. No. 2.3e+03;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 9 SKDFPNMSETDL 20
 :||| :||| :
 DB 2 SKDFHIVAEETGI 13

Search completed: July 15, 2004, 14:51:01
 Job time : 27.7333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:47:02 ; Search time 12.4 Seconds
(without alignments)
83.268 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101

Sequence: 1 DLEETAAANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 204689

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	32.7	20	6	5210075-6
2	32	31.7	20	4	US-09-161-939A-9
3	32	31.7	23	3	US-08-847-844A-69
4	30	29.7	16	1	US-08-446-692-53
5	30	29.7	16	2	US-08-488-351A-53
6	30	29.7	16	3	US-09-100-414B-78
7	30	29.7	16	3	US-09-100-409A-7
8	30	29.7	16	3	US-09-303-323-78
9	30	29.7	16	4	US-09-770-014-78
10	30	29.7	16	5	PCT-US95-13841-3
11	30	29.7	16	1	US-08-080-073-27
12	29	28.7	15	1	US-10-053-485-14
13	29	28.7	16	4	US-08-218-025A-54
14	29	28.7	20	1	US-09-161-939A-1
15	29	28.7	20	4	US-08-232-513A-7
16	28	27.7	15	3	US-09-231-159-12
17	28	27.7	15	3	US-08-611-307-12
18	28	27.7	26	1	US-08-446-692-84
19	28	27.7	26	2	US-08-488-351A-84
20	28	27.7	26	4	US-09-765-815-5
21	28	27.7	28	3	US-08-416-711-4
22	28	27.7	28	4	US-09-356-497-4
23	28	27.7	27	2	US-08-563-892A-3
24	27.5	27.2	27	4	US-09-366-212A-3
25	27.5	27.2	12	4	US-09-025-596-21
26	27	26.7	12	4	US-08-874-197-4
27	27	26.7	15	3	

28 27 26.7 15 4 US-08-648-182-4 Sequence 4, Appli
29 27 26.7 15 4 US-09-387-671-38 Sequence 38, Appl
30 27 26.7 15 4 US-09-387-671-43 Sequence 43, Appl
31 27 26.7 16 4 US-09-387-671-37 Sequence 37, Appl
32 27 26.7 16 4 US-09-387-671-42 Sequence 42, Appl
33 27 26.7 17 4 US-09-387-671-36 Sequence 36, Appl
34 27 26.7 17 4 US-09-387-671-41 Sequence 41, Appl
35 27 26.7 18 4 US-09-387-671-35 Sequence 35, Appl
36 27 26.7 18 4 US-09-387-671-40 Sequence 40, Appl
37 27 26.7 19 4 US-09-387-671-34 Sequence 34, Appl
38 27 26.7 19 4 US-09-387-671-39 Sequence 39, Appl
39 27 26.7 20 4 US-09-687-875A-22 Sequence 22, Appl
40 27 26.7 20 4 US-09-387-671-44 Sequence 44, Appl
41 27 26.7 20 4 US-09-387-671-45 Sequence 45, Appl
42 27 26.7 20 4 US-09-387-671-46 Sequence 46, Appl
43 27 26.7 20 4 US-09-387-671-47 Sequence 47, Appl
44 27 26.7 20 4 US-09-387-671-50 Sequence 50, Appl
45 27 26.7 20 4

ALIGNMENTS

RESULT 1
5210075-6
; Patent No. 5210075
; APPLICANT: SCHOLTZ, WOLFGANG; CHIANG, SHIU-LANG; NAGARAJAN,
; GOBI; LOBL, THOMAS J.
; TITLE OF INVENTION: INTERLEUKIN 6 ANTAGONIST PEPTIDES
; NUMBER OF SEQUENCES: 69
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,868
; FILING DATE: 16-FEB-1990
; SEQ ID NO: 6:
; LENGTH: 20
5210075-6

Query Match 32.7%; Score 33; DB 6; Length 20;
Best Local Similarity 40.0%; Pred. No. 39;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 BEIAANSKDFPNMSE 17
DB 5 EALAENLNLPKMAE 19

RESULT 2
US-09-161-939A-9
; Sequence 9, Application US/09161939A
; Patent No. 6486299
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Genes and Proteins Predictive and Therapeutic for
; FILE REFERENCE: 15966-527
; CURRENT APPLICATION NUMBER: US/09/161,939A
; CURRENT FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-161-939A-9

Query Match 31.7%; Score 32; DB 4; Length 20;
Best Local Similarity 40.0%; Pred. No. 57;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 BEIAANSKDFPNMSE 17
DB 1 EEVAAATRRLEDISE 15

RESULT 3

US-08-847-844A-69
 ; Sequence 69, Application US/08847844A
 ; Patent No. 6150160
 ; GENERAL INFORMATION:
 ; APPLICANT: KAZAZIAN JR., HAIG H.
 ; APPLICANT: BOEKE, JEFF D.
 ; APPLICANT: MORAN, JOHN V.
 ; APPLICANT: DOMBROSKI, BETH A.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE OF
 ; MAMMALIAN RETROTRANSPOSONS
 ; NUMBER OF SEQUENCES: 137
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PANITCH SCHWARZ JACOBS & NADEL, P.C.
 ; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND FL.
 ; CITY: PHILADELPHIA
 ; STATE: PA
 ; COUNTRY: U.S.A.
 ; ZIP: 19103-7086
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/847,844A
 ; FILING DATE: 28-APR-1997
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/749,805
 ; FILING DATE: 16-NOV-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/006,831
 ; FILING DATE: 16-NOV-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DOYLE LEARY PH.D., KATHRYN
 ; REGISTRATION NUMBER: 36,317
 ; REFERENCE/DOCKET NUMBER: 9596-2302
 ; TELEPHONE: 215-567-2020
 ; TELEFAX: 215-567-2991
 ; INFORMATION FOR SEQ ID NO: 69:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 23 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-847-844A-69

Query Match 31.7%; Score 32; DB 3; Length 23;
 Best Local Similarity 46.2%; Pred. No. 68;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy

7 ANSKDFPNMSETD 19

Db

6 SNEQDDFTMSRID 18

RESULT 4

US-08-446-692-53
 ; Sequence 53, Application US/08446692
 ; Patent No. 5759551
 ; GENERAL INFORMATION:
 ; APPLICANT: Ladd, Anna
 ; APPLICANT: Wang, Chang Yi
 ; APPLICANT: Zamb, Timothy
 ; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
 ; AND SYNTHETIC UNIVERSAL IMMUNE STIMULATORS FOR VACCINES
 ; NUMBER OF SEQUENCES: 114
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Maria C.H. Lin

STREET: 345 Park Avenue
 CITY: New York
 STATE: NY
 COUNTRY: US
 ZIP: 10154-0053
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/446,692
 ; FILING DATE: 7-JUN-1995
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maria C.H. Lin
 ; REGISTRATION NUMBER: 29,323
 ; REFERENCE/DOCKET NUMBER: 1151-4146 US2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212)415-8745
 ; TELEFAX: (516)751-6849
 ; INFORMATION FOR SEQ ID NO: 53:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 16 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-446-692-53

Query Match 29.7%; Score 30; DB 1; Length 16;
 Best Local Similarity 50.0%; Pred. No. 95;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy

7 ANSKDFPNMSET 18

Db

2 AKSKKFPSTAT 13

RESULT 5

US-08-488-351A-53
 ; Sequence 53, Application US/08488351A
 ; Patent No. 5843446
 ; GENERAL INFORMATION:
 ; APPLICANT: Ladd, Anna
 ; APPLICANT: Wang, Chang Yi
 ; APPLICANT: Zamb, Timothy
 ; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
 ; AND SYNTHETIC UNIVERSAL IMMUNE STIMULATORS FOR VACCINES
 ; NUMBER OF SEQUENCES: 114
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Maria C.H. Lin
 ; STREET: 345 Park Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: US
 ; ZIP: 10154-0053
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/488,351A
 ; FILING DATE: 7-JUN-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/446,692
 ; FILING DATE: 7-JUN-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/229,275
 ; FILING DATE: 14-APR-1994
 ; CLASSIFICATION: 424

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-488-351A-53

Query Match 29.7%; Score 30; DB 2; Length 16;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 ANSKDFPNMSET 18
Db 2 AKSKPPSYTAT 13

RESULT 6
US-09-100-414B-78
; Sequence 78, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL IHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-78

Query Match 29.7%; Score 30; DB 3; Length 16;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 ANSKDFPNMSET 18

Db 2 AKSKPPSYTAT 13

RESULT 7
US-09-100-409A-7
; Sequence 7, Application US/09100409A
; Patent No. 6090388
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: PEPTIDE COMPOSITION FOR
; PREVENTION AND TREATMENT OF HIV INFECTION AND
; IMMUNE DISORDERS
; TITLE OF INVENTION: IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,409A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 1151-4154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-409A-7

Query Match 29.7%; Score 30; DB 3; Length 16;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 ANSKDFPNMSET 18
Db 2 AKSKPPSYTAT 13

RESULT 8
US-09-100-600A-22
; Sequence 22, Application US/09100600A
; Patent No. 6107021
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Shen, Ming
; TITLE OF INVENTION: SYNTHETIC PEPTIDE VACCINES FOR
; FOOT-AND-MOUTH DISEASE
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
FILING DATE: 20-Jun-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4156
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-600A-22

Query Match 29.7%; Score 30; DB 3; Length 16;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 ANSKDFPNMSET 18
DB 2 AKSKKFPSTAT 13

RESULT 9

US-09-303-323-78
; Sequence 78, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; FILING DATE: 30-APR-1999
; APPLICATION NUMBER: US/09/303,323
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-09-303-323-78

Query Match 29.7%; Score 30; DB 3; Length 16;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 ANSKDFPNMSET 18
DB 2 AKSKKFPSTAT 13

RESULT 10

US-09-770-014-78
; Sequence 78, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-78

Query Match 29.7%; Score 30; DB 4; Length 16;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 ANSKDFPNMSET 18
DB 2 AKSKKFPSTAT 13

RESULT 11

PCT-US95-13841-3
; Sequence 3, Application PC/TUS9513841
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc; Walfield, Alan M.;
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Synthetic IGE Membrane Anchor
; TITLE OF INVENTION: Peptide Immunogens for the Treatment of Allergy
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:


```
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-3

Query Match 29.7%; Score 30; DB 5; Length 16;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 ANSKDFPNMSET 18
Db 2 AKSKKFPSTAT 13

RESULT 12
US-08-080-073-27
; Sequence 27, Application US/08080073
; Patent No. 5384255
; GENERAL INFORMATION:
; APPLICANT: Ciechanover, Aaron J.
; APPLICANT: Blumenfeld, Nava
; APPLICANT: Gonen, Hedva
; TITLE OF INVENTION: Ubiquitin Carrier Enzyme E2-F1,
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/080,073
; FILING DATE: 21-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldstein, Jorge A.

; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13841
; FILING DATE: 25-OCT-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/328,519
; FILING DATE: 25-OCT-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin, Maria C.H.
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4117
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13841-3

Query Match 29.7%; Score 29; DB 4; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9 SKDFPNMSETDL 20
Db 2 STDFFNFDGSL 13

RESULT 14
US-08-218-025A-54
; Sequence 54, Application US/08218025A
; Patent No. 5556744
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B.
; APPLICANT: Ugen, Kenneth E.
; APPLICANT: Williams, William V.
; TITLE OF INVENTION: Methods and Compositions for Diagnosing
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457, 321 No. 555674Aristown Road
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1448.0050000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-080-073-27

Query Match 28.7%; Score 29; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 SKDFPNMSET 18
Db 6 SEDYPNKPT 15

RESULT 13
US-10-053-485-14
; Sequence 14, Application US/10053485
; Patent No. 6576896
; GENERAL INFORMATION:
; APPLICANT: Figeys, Daniel
; APPLICANT: Aebersold, Ruedi
; TITLE OF INVENTION: ELECTROSMOTIC FLUIDIC DEVICE AND RELATED METHODS
; FILE REFERENCE: UWOPL118617
; CURRENT APPLICATION NUMBER: US/10/053,485
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: US 09/209,880
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 60/069,398
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 14
; LENGTH: 16
; TYPE: PPT
; ORGANISM: Saccharomyces cerevisiae
US-10-053-485-14

Query Match 28.7%; Score 29; DB 4; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9 SKDFPNMSETDL 20
Db 2 STDFFNFDGSL 13

RESULT 14
US-08-218-025A-54
; Sequence 54, Application US/08218025A
; Patent No. 5556744
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B.
; APPLICANT: Ugen, Kenneth E.
; APPLICANT: Williams, William V.
; TITLE OF INVENTION: Methods and Compositions for Diagnosing
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457, 321 No. 555674Aristown Road
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.2.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,025A
FILING DATE: 24-MAR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/891,451
FILING DATE: 29-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST33A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-218-025A-54

```

```

Query Match      28.7%; Score 29; DB 1; Length 20;
Best Local Similarity 31.2%; Pred. No. 1.8e+0e;
Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 3 EETAANSKQPPNMSET 18
    ||: ||: : ||
Db 3 EEVVRSENFDTNAKT 18

```

```

RESULT 15
US-09-161-939A-1
; Sequence 1, Application US/09161939A
; Patent No. 6486299
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: Genes and Proteins Predictive and Therapeutic for
; Stroke, Hypertension, Diabetes, and Obesity
; FILE REFERENCE: 15966-527
; CURRENT APPLICATION NUMBER: US/09/161,939A
; CURRENT FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-161-939A-1

```

```

Query Match      28.7%; Score 29; DB 4; Length 20;
Best Local Similarity 40.0%;
Pred. No. 1.8e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QV      3 EEIAANSKQFPNMSE 17
Db      1 EEVAATTRROEDISE 15

```

Search completed: July 15, 2004, 15:13:21
Job time : 12.4 secs

Result No.	Score	Query %			ID	Description
		Match	Length	DB		
1	101	100.0	20	10	US-09-171-432A-45	Sequence 45, Appl
2	53	52.5	20	10	US-09-171-432A-44	Sequence 44, Appl
3	36	35.6	10	12	US-10-462-452-709	Sequence 709, Appl
4	36	35.6	10	16	US-10-601-953-814	Sequence 814, Appl
5	36	35.6	10	16	US-10-322-866-710	Sequence 710, Appl
6	35	34.7	10	16	US-10-447-161-80	Sequence 80, Appl
7	33	32.7	15	14	US-10-175-270-12	Sequence 12, Appl
8	33	32.7	25	16	US-10-416-249-78	Sequence 78, Appl
9	32	31.7	15	12	US-10-346-162-187	Sequence 187, Appl
10	32	31.7	15	12	US-10-346-162-239	Sequence 239, Appl
11	32	31.7	20	14	US-10-196-703-9	Sequence 9, Appl
12	32	31.7	23	10	US-09-927-734C-5	Sequence 5, Appl
13	32	31.7	23	14	US-10-216-122-69	Sequence 69, Appl
14	31	30.7	14	12	US-10-375-157-24	Sequence 24, Appl
15	31	30.7	14	14	US-10-010-568-24	Sequence 24, Appl

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label=YK-1328
US-09-171-432A-45

Query Match 100.0%; Score 101; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAASKDPNNSETDL 20
Db 1 DLEETAAASKDPNNSETDL 20

RESULT 2
US-09-171-432A-44

; Sequence 44, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudiyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label=YK-1327
US-09-171-432A-44

Query Match 52.5%; Score 53; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.094;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKD 11
Db 10 DLEETAAANSKD 20

RESULT 3

US-10-462-452-709
; Sequence 709, Application US/10462452
; Publication No. US20040037809A1
; GENERAL INFORMATION:
; APPLICANT: Quay, Steven
; APPLICANT: El Shafy, Mohammed Abd
; APPLICANT: Gupta, Malini
; APPLICANT: de Meireles, Jorge
; TITLE OF INVENTION: Compositions and Methods for Enhanced
; TITLE OF INVENTION: Mucosal-Delivery of Interferon Beta
; FILE REFERENCE: 02-02US
; CURRENT APPLICATION NUMBER: US/10/462,452
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: 60/393,066
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 790
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 709
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-462-452-709

Query Match 35.6%; Score 36; DB 12; Length 10;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ANSKDFPNM 15
Db 1 SNCKDFPSM 9

RESULT 4

US-10-601-953-814
; Sequence 814, Application US/10601953
; Publication No. US20040077540A1
; GENERAL INFORMATION:
; APPLICANT: Quay, Steven C.
; TITLE OF INVENTION: Compositions And Methods For Modulating Physiology Of Epithelial
; TITLE OF INVENTION: Junctional Adhesion Molecules For Enhanced Mucosal Delivery Of
; TITLE OF INVENTION: Therapeutic Compounds
; FILE REFERENCE: 02-03US
; CURRENT APPLICATION NUMBER: US/10/601,953
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: 60/392,512
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 814
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-10-601-953-814

Query Match 35.6%; Score 36; DB 16; Length 10;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ANSKDFPNM 15
Db 1 SNCKDFPSM 9

RESULT 5

US-10-322-266-710

```
; Sequence 710, Application US/10322266
; Publication No. US2004011535A1
; GENERAL INFORMATION:
; APPLICANT: Quay, Steven C.
; TITLE OF INVENTION: Compositions And Methods For Enhanced Mucosal Delivery Of Peptide
; FILE REFERENCE: YV And Methods For Treating And Preventing Obesity
; CURRENT APPLICATION NUMBER: US/10/322,266
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 797
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 710
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-10-322-266-710

Query Match          35.6%; Score 36; DB 16; Length 10;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ANSKDPPNM 15
Db 1 SNCKDPPSM 9

RESULT 6
US-10-447-161-80
; Sequence 80, Application US/10447161
; Publication No. US20040023314A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis
; FILE REFERENCE: HO-P02484US1
; CURRENT APPLICATION NUMBER: US/10/447,161
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/383,530
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-447-161-80

Query Match          34.7%; Score 35; DB 16; Length 10;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 11 DFPNMSETDL 20
Db 1 DYPNLSATDI 10

RESULT 7
US-10-175-270-12
; Sequence 12, Application US/10175270
; Publication No. US20030166548A1
; GENERAL INFORMATION:
; APPLICANT: PETERSON, JOHNNY W.
; APPLICANT: SAINTI, SHAMSHER S.
; APPLICANT: WOOD, THOMAS G.
; APPLICANT: CHOPRA, ASHOK K.
; TITLE OF INVENTION: ANTI-INFLAMMATORY AND OTHER THERAPEUTIC PROPHYLACTIC OR DIAGNOSTIC
; TITLE OF INVENTION: SYNTHETIC MELITTIN AND NE
; TITLE OF INVENTION: W RELATED PEPTIDES
; FILE REFERENCE: UTSG:213
; CURRENT APPLICATION NUMBER: US/10/175,270
```

```
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/458,397
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: PCT/US98/12340
; PRIOR FILING DATE: 1998-06-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-175-270-12

Query Match          32.7%; Score 33; DB 14; Length 15;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 NSKDFPNMSE 17
Db 1 NSKDFVTSE 10

RESULT 8
US-10-416-249-78
; Sequence 78, Application US/10416249
; Publication No. US20040132121A1
; GENERAL INFORMATION:
; APPLICANT: Dalrymple, Brian P.
; APPLICANT: Kongsuwan, Kritaya
; APPLICANT: Wijffels, Gene L.
; APPLICANT: Jennings, Philip A.
; APPLICANT: Kemp, Gregory W.
; TITLE OF INVENTION: METHOD OF IDENTIFYING ANTIBACTERIAL
; FILE REFERENCE: CULIN42.001APC
; CURRENT APPLICATION NUMBER: US/10/416,249
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: PCT/AU01/01436
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: AU PR 1320
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: AU PR 2919
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 678
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-416-249-78

Query Match          32.7%; Score 33; DB 16; Length 25;
Best Local Similarity 36.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 LEEIAANSKDFPNMSETDL 20
Db 2 LDAQIADAETFPFMGERRL 20

RESULT 9
US-10-346-162-187
; Sequence 187, Application US/10346162
; Publication No. US20030224390A1
; GENERAL INFORMATION:
; APPLICANT: KARO BIO USA, INC.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: BARNETT, Thomas R.
; APPLICANT: BUEHRER, Benjamin
; TITLE OF INVENTION: METHOD OF IDENTIFYING CONFORMATION-SENSITIVE BINDING PEPTIDES AND
```

;; TITLE OF INVENTION: THEREOF
;; FILE REFERENCE: PAIGE-1H
;; CURRENT APPLICATION NUMBER: US/10/346,162
;; CURRENT FILING DATE: 2003-01-17
;; PRIOR APPLICATION NUMBER: US 09/614,865
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 09/860,688
;; PRIOR FILING DATE: 2001-05-21
;; NUMBER OF SEQ ID NOS: 268
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 187
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: synthetic
US-10-346-162-187

Query Match 31.7%; Score 32; DB 12; Length 15;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KDFPNM 15
|||
Db 4 KDFPNL 9

RESULT 10

US-10-346-162-239
;; Sequence 239, Application US/10346162
;; Publication No. US20030224390A1
;; GENERAL INFORMATION:
;; APPLICANT: KARO BIO USA, INC.
;; APPLICANT: FOMLKEs, Dana M.
;; APPLICANT: BUHRER, Benjamin
;; TITLE OF INVENTION: METHOD OF IDENTIFYING CONFORMATION-SENSITIVE BINDING PEPTIDES AND
;; TITLE REFERENCE: PAIGE-1H
;; CURRENT APPLICATION NUMBER: US/10/346,162
;; CURRENT FILING DATE: 2003-01-17
;; PRIOR APPLICATION NUMBER: US 09/614,865
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 09/860,688
;; PRIOR FILING DATE: 2001-05-21
;; NUMBER OF SEQ ID NOS: 268
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 239
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: synthetic
US-10-346-162-239

Query Match 31.7%; Score 32; DB 12; Length 15;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KDFPNM 15
|||
Db 4 KDFPNL 9

RESULT 11

US-10-196-703-9
;; Sequence 9, Application US/10196703
;; Publication No. US20030055019A1
;; GENERAL INFORMATION:
;; APPLICANT: Shimkets, Richard A.
;; TITLE OF INVENTION: Genes and Proteins Predictive and Therapeutic for
;; TITLE OF INVENTION: Stroke, Hypertension, Diabetes, and Obesity
;; FILE REFERENCE: 15966-527

;; CURRENT APPLICATION NUMBER: US/10/196,703
;; CURRENT FILING DATE: 2002-07-15
;; PRIOR APPLICATION NUMBER: US/09/161,939
;; PRIOR FILING DATE: 1998-09-28
;; NUMBER OF SEQ ID NOS: 44
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 9
;; LENGTH: 20
;; TYPE: PRT
;; ORGANISM: Rattus sp.
US-10-196-703-9

Query Match 31.7%; Score 32; DB 14; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 3 EEIAANSKDFENMGE 17
|||
Db 1 EEVAATRRLEDEISE 15

RESULT 12

US-09-927-734C-5
;; Sequence 5, Application US/09927734C
;; Publication No. US20030194758A1
;; GENERAL INFORMATION:
;; APPLICANT: Selltreinnikoff, Claude
;; APPLICANT: Nakata, Mitsunori
;; TITLE OF INVENTION: Method for the Identification of Fungal Glucose Utilization
;; FILE REFERENCE: MYCOLOGX-06279
;; CURRENT APPLICATION NUMBER: US/09/927,734C
;; CURRENT FILING DATE: 2001-08-10
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 5
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Saccharomyces cerevisiae
US-09-927-734C-5

Query Match 31.7%; Score 32; DB 10; Length 23;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LEEIAANSK 10
|||
Db 1 IEDIAVNSK 9

RESULT 13

US-10-216-122-69
;; Sequence 69, Application US/10216122
;; Publication No. US20030121063A1
;; GENERAL INFORMATION:
;; APPLICANT: Kazazian, Haig H.
;; APPLICANT: Osterlag, Eric
;; APPLICANT: Debernardinis, Ralph
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE OF MAMMALIAN RETROTRANSPOSONS
;; FILE REFERENCE: 053893-5006-03
;; CURRENT APPLICATION NUMBER: US/10/216,122
;; CURRENT FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: US 09/653,812
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: US 08/847,844
;; PRIOR FILING DATE: 1997-04-28
;; PRIOR APPLICATION NUMBER: US 08/749,805
;; PRIOR FILING DATE: 1996-11-15
;; PRIOR APPLICATION NUMBER: US 60/006,831
;; PRIOR FILING DATE: 1995-11-16
;; NUMBER OF SEQ ID NOS: 154
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 69

; LENGTH: 23		Matches		5;	Conservative	2;	Mismatches	1;	Indels	0;	Gaps	0;											
; TYPE: PRT																							
; ORGANISM: Zea mays																							
US-10-216-122-69																							
Query Match		31.7%;		Score 32;		DB 14;		Length 23;															
Best Local Similarity		46.2%;		Pred. No. 3.2e+02;																			
Matches		6;	Conservative	2;	Mismatches	5;	Indels	0;	Gaps	0;													
QY		7	ANSKDFPNMSETD 19																				
Db		6	SNEQDDPTMSRID 18																				
RESULT 14																							
US-10-375-157-24																							
; Sequence 24,		Application US/10375157																					
; Publication No.		US20030224458A1																					
; GENERAL INFORMATION:																							
; APPLICANT:		Bristol-Myers Squibb Company																					
; TITLE OF INVENTION:		A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY23, EXPRESSED																					
; FILE REFERENCE:		D0077A CIP																					
; CURRENT APPLICATION NUMBER:		US/10/375,157																					
; PRIOR FILING DATE:		2003-02-26																					
; PRIOR APPLICATION NUMBER:		US 60/251,926																					
; PRIOR FILING DATE:		2000-12-07																					
; PRIOR APPLICATION NUMBER:		US 10/010,568																					
; PRIOR FILING DATE:		2001-12-07																					
; PRIOR APPLICATION NUMBER:		US 60/269,795																					
; PRIOR FILING DATE:		2001-02-14																					
; NUMBER OF SEQ ID NOS:		65																					
; SOFTWARE:		PatentIn version 3.2																					
; SEQ ID NO 24																							
; LENGTH: 14																							
; TYPE: PRT																							
; ORGANISM: homo sapiens																							
US-10-375-157-24																							
Query Match		30.7%;		Score 31;		DB 12;		Length 14;															
Best Local Similarity		62.5%;		Pred. No. 2.6e+02;																			
Matches		5;	Conservative	2;	Mismatches	1;	Indels	0;	Gaps	0;													
QY		7	ANSKDFPN 14																				
Db		5	ANASDFPD 12																				
RESULT 15																							
US-10-010-568-24																							
; Sequence 24,		Application US/10010568																					
; Publication No.		US20030157598A1																					
; GENERAL INFORMATION:																							
; APPLICANT:		Bristol-Myers Squibb Company																					
; TITLE OF INVENTION:		A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY23, EXPRESSED HI																					
; FILE REFERENCE:		D0077 NP																					
; CURRENT APPLICATION NUMBER:		US/10/010,568																					
; PRIOR FILING DATE:		2001-12-07																					
; PRIOR APPLICATION NUMBER:		US 60/251,926																					
; PRIOR FILING DATE:		2000-12-07																					
; PRIOR APPLICATION NUMBER:		US 60/269,795																					
; PRIOR FILING DATE:		2001-02-14																					
; NUMBER OF SEQ ID NOS:		55																					
; SOFTWARE:		PatentIn version 3.0																					
; SEQ ID NO 24																							
; LENGTH: 14																							
; TYPE: PRT																							
; ORGANISM: homo sapiens																							
US-10-010-568-24																							
Query Match		30.7%;		Score 31;		DB 14;		Length 14;															
Best Local Similarity		62.5%;		Pred. No. 2.6e+02;																			

Search completed: July 15, 2004, 15:22:57
Job time : 33.5333 secs

[Faint, illegible text]
THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:36:31 ; Search time 40.1333 Seconds
(without alignments)
140.804 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRMGLSGVQBIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 581562

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_29Jan04.*
1: Genesep1980s.*
2: Genesep1990s.*
3: Genesep2000s.*
4: Genesep2001s.*
5: Genesep2002s.*
6: Genesep2003as.*
7: Genesep2003bs.*
8: Genesep2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	20	2 AAW42929	Immunogen
2	96	100.0	20	4 AAB69446	Synthetic
3	96	100.0	25	2 AAW42969	Immunogen
4	96	100.0	25	4 AAB69448	Synthetic
5	35	36.5	14	5 AAG96117	Synthetic
6	33	34.4	27	2 AAY07977	Cysteine-
7	32	33.3	10	3 AAB27118	Human sec
8	32	33.3	20	4 AAB81167	Human CAS
9	32	33.3	20	6 AABU07394	Foreign rel
10	32	33.3	23	7 ADD13104	Foreign p
11	32	33.3	25	5 AAG62321	Human IDH
12	32	33.3	28	2 AAW77325	Eubacteri
13	31	32.3	9	3 AAB26182	Human Rep
14	31	32.3	14	5 AAO18055	Human neu
15	31	32.3	14	5 AAO18057	Human neu
16	31	32.3	15	4 AAO00745	Human hta
17	31	32.3	19	5 ABB82306	Epitopic
18	31	32.3	20	5 AAG79500	Epitopic
19	31	32.3	20	6 AAE29341	Human Ban
20	31	32.3	20	6 AAE38127	Human cyt
21	31	32.3	23	4 AAE05961	Signal pe
22	31	32.3	25	2 AAR88164	Leader se
23	30	31.2	13	4 ABR51762	Mutant Pe
24	30	31.2	19	2 AAR86351	HIV-1 pro
25	30	31.2	20	3 AAB33569	5/20 nati

26	30	31.2	20	3 AAB23120	Peanut wi
27	30	31.2	20	4 AAU04782	Native an
28	30	31.2	20	4 AAU05107	Anaphylac
29	30	31.2	21	6 ABP72310	Plasmodiu
30	30	31.2	21	7 ADE79947	Plasmodiu
31	30	31.2	23	6 ADB16892	Xylella f
32	30	31.2	25	5 ABG62324	Eubacteri
33	30	31.2	25	5 ABG62323	Eubacteri
34	30	31.2	28	2 AAW62078	Eubacteri
35	29	30.2	13	4 ABR51694	Hyphozyma
36	29	30.2	19	1 AAP91675	Mutant Pe
37	29	30.2	19	2 AAR86352	Heavy cha
38	29	30.2	19	5 AAG96214	HIV-1 pro
39	29	30.2	20	2 AAW42907	Cysteine-
40	29	30.2	21	4 AAB69424	Immunogen
41	29	30.2	21	6 AAB32173	Synthetic
42	29	30.2	22	2 AAY18852	Haemophil
43	29	30.2	22	2 AAY18852	Lecithin:
44	29	30.2	22	2 AAY19106	Lecithin:
45	29	30.2	22	2 AAY19360	Lecithin:
	29	30.2	22	2 AAY18589	Lecithin:

ALIGNMENTS

RESULT 1
AAW42929
ID AAW42929 standard; peptide; 20 AA.
XX
AC AAW42929;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1331.
KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006891.
XX
PR 19-APR-1996; 96US-0015644P.
XX
EA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
FI Fields HA, Khudyakov YE;
XX
DR WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 961-980. The present peptide
CC is derived from amino acids 961-980, and has a reactivity of 27.1% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
XX administering the peptide to a mammal
SQ Sequence 20 AA;

Fri Jul 16 09:25:19 2004

Query Match 100.0%; Score 96; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 DB 1 KINLADRMGLSGVQEIKEQ 20

RESULT 2
 AAB69446
 ID AAB69446 standard; peptide; 20 AA.
 XX
 AC AAB69446;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 46.
 XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 XX
 OS Hepatitis A virus.
 OS Synthetic.
 PN WO200105824-A2.
 XX
 PD 25-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-US019267.
 PR 15-JUL-1999; 99US-0144412P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields HA, Khudyakov YE;
 XX
 PF IPI; 2001-112681/12.
 XX
 PT Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-Hepatitis A virus and as vaccines.
 XX
 PS Claim 13; Page 97; 130pp; English.

The present sequence is one of a number of synthetic peptides which are
 immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 comprise antigenic epitopes of the major structural capsid polypeptides
 or non-structural polypeptides of HAV with one or more glutamine
 molecules at the carboxy end of the peptide. The peptides are used to
 detect the presence of antibodies against HAV in mammalian serum, to
 detect the presence of HAV in a human or animal through the binding of
 the peptide to an antibody, to detect acute phase infection by detecting
 IGM antibodies in mammalian serum and detecting convalescence in a
 mammal. The peptides are used to detect or quantify HAV antibodies in
 samples in clinical or research-based assays using immunoblotting,
 fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 tracking of radioactive or bioluminescent markers, chromatography or
 electrophoresis. The peptides are used to induce an immune response to
 HAV when administered to a human or animal. Glutamine at the carboxy end
 of the peptides enhances the IGM antibody reactivity

Sequence 20 AA;
 Query Match 100.0%; Score 96; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 DB 1 KINLADRMGLSGVQEIKEQ 20

RESULT 3
 AAW42969
 ID AAW42969 standard; peptide; 25 AA.
 XX
 AC AAW42969;
 XX
 DT 28-APR-1998 (first entry)
 XX
 DE Immunogenic Hepatitis A virus peptide YK-1757.
 XX
 KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
 KW antibody.
 OS Synthetic.
 OS Hepatitis A virus.
 XX
 PN WO9740147-A1.
 XX
 PD 30-OCT-1997.
 XX
 PF 18-APR-1997; 97WO-US006891.
 XX
 PR 19-APR-1996; 96US-0015644P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields HA, Khudyakov YE;
 XX
 PF IPI; 1997-535831/49.
 XX
 PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.
 XX
 PS Claim 18; Page 112; 140pp; English.

The present immunogenic peptide corresponds to an immunogenic epitope of
 the Hepatitis A virus (HAV). The peptide is substantially similar to a
 portion of the amino acid sequence of the P2A protein of HAV
 corresponding to amino acids 792-980. Compositions containing the peptide
 can be used to induce an immune response to HAV in a mammal. The peptide
 can also be used to detect the presence of antibodies against HAV in
 mammalian serum. The peptide can also be used to make an antibody against
 HAV by administering the peptide to a mammal

Sequence 25 AA;
 Query Match 100.0%; Score 96; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 DB 6 KINLADRMGLSGVQEIKEQ 25

RESULT 4
 AAB69448
 ID AAB69448 standard; peptide; 25 AA.
 XX
 AC AAB69448;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 48.
 XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 XX
 OS Hepatitis A virus.
 OS Synthetic.
 PN WO200105824-A2.

Query Match 100.0%; Score 96; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 DB 1 KINLADRMGLSGVQEIKEQ 20

RESULT 2
 AAB69446
 ID AAB69446 standard; peptide; 20 AA.
 XX
 AC AAB69446;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 46.
 XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 XX
 OS Hepatitis A virus.
 OS Synthetic.
 PN WO200105824-A2.
 XX
 PD 25-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-US019267.
 PR 15-JUL-1999; 99US-0144412P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields HA, Khudyakov YE;
 XX
 PF IPI; 2001-112681/12.
 XX
 PT Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-Hepatitis A virus and as vaccines.
 XX
 PS Claim 13; Page 97; 130pp; English.

The present sequence is one of a number of synthetic peptides which are
 immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 comprise antigenic epitopes of the major structural capsid polypeptides
 or non-structural polypeptides of HAV with one or more glutamine
 molecules at the carboxy end of the peptide. The peptides are used to
 detect the presence of antibodies against HAV in mammalian serum, to
 detect the presence of HAV in a human or animal through the binding of
 the peptide to an antibody, to detect acute phase infection by detecting
 IGM antibodies in mammalian serum and detecting convalescence in a
 mammal. The peptides are used to detect or quantify HAV antibodies in
 samples in clinical or research-based assays using immunoblotting,
 fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 tracking of radioactive or bioluminescent markers, chromatography or
 electrophoresis. The peptides are used to induce an immune response to
 HAV when administered to a human or animal. Glutamine at the carboxy end
 of the peptides enhances the IGM antibody reactivity

Sequence 20 AA;
 Query Match 100.0%; Score 96; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 DB 1 KINLADRMGLSGVQEIKEQ 20

Query Match 100.0%; Score 96; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
 DB 6 KINLADRMGLSGVQEIKEQ 25

RESULT 4
 AAB69448
 ID AAB69448 standard; peptide; 25 AA.
 XX
 AC AAB69448;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 48.
 XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 XX
 OS Hepatitis A virus.
 OS Synthetic.
 PN WO200105824-A2.

```

XX PD 25-JAN-2001.
XX XX
XX PF 14-JUL-2000; 2000WO-US019267.
XX XX
XX PR 15-JUL-1999; 99US-0144412P.
XX XX
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX XX
XX PI Fields HA, Khudyakov YE;
XX XX
XX DR WPI; 2001-112681/12.
XX XX
XX PT Synthetic peptides used as antigen sources for enzyme immunoassays
XX PT detecting anti-hepatitis A virus and as vaccines.
XX PS
XX PS Claim 13; Page 99; 130pp; English.
XX XX
XX CC The present sequence is one of a number of synthetic peptides which are
XX CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
XX CC comprise antigenic epitopes of the major structural capsid polypeptides
XX CC or non-structural polypeptides of HAV with one or more glutamine
XX CC molecules at the carboxy end of the peptide. The peptides are used to
XX CC detect the presence of antibodies against HAV in mammalian serum, to
XX CC detect the presence of HAV in a human or animal through the binding of
XX CC the peptide to an antibody, to detect acute phase infection by detecting
XX CC IgM antibodies in mammalian serum and detecting convalescence in a
XX CC mammal. The peptides are used to detect or quantify HAV antibodies in
XX CC samples in clinical or research-based assays using immunoblotting,
XX CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
XX CC tracking of radioactive or bioluminescent markers, chromatography or
XX CC electrophoresis. The peptides are used to induce an immune response to
XX CC HAV when administered to a human or animal. Glutamine at the carboxy end
XX CC of the peptides enhances the IgM antibody reactivity
XX XX
XX SQ Sequence 25 AA;
XX
XX Query Match 100.0%; Score 96; DB 4; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-09;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KINLADRLGLSGVQEIKEQ 20
XX Db ||||| ||||| ||||| |||||
XX 6 KINLADRLGLSGVQEIKEQ 25
XX
XX RESULT 5
XX ABG96117
XX ID ABG96117 standard; peptide; 14 AA.
XX AC ABG96117;
XX XX
XX DT 11-DEC-2002 (first entry)
XX XX
XX DE Cysteine-containing peptide isolated by the invention method, #90.
XX KW Rabbit; bovine; analytical reagent; trifuunctional; peptide mixture;
XX KW enrichment; immobilisation site; cleavage site; link; epitope tag;
XX KW protease; cysteine-containing; perturbed cell; mass spectrometry;
XX KW peptide tag; BSA; bovine serum albumin; PEPTag; APEPTag; IPEPTag;
XX KW affinity peptide encoded tag; immobilised peptide encoded tag; chicken;
XX KW beta-lactoglobulin; GAPDH; glyceraldehyde-3-phosphate dehydrogenase;
XX KW a-lactalbumin; ovalbumin; yeast.
XX XX
XX OS Saccharomyces cerevisiae.
XX XX
XX PN WO200259144-A2.
XX XX
XX PD 01-AUG-2002.
XX XX
XX PF 25-JAN-2002; 2002WO-US002487.
XX XX
XX PR 26-JAN-2001; 2001US-0264576P.
XX XX

```

```

PR 13-JUL-2001; 2001US-0305232P.
XX XX
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX XX
XX PI Haynes P, Wei J, Yates J, Andon N;
XX XX
XX DR WPI; 2002-599760/64.
XX XX
XX PT Novel trifunctional synthetic reagents for labeling peptides at specific
XX PT amino acid residue and selectively enriching only those peptides
XX PT containing labeled amino acid, useful for proteomic analysis.
XX XX
XX PS Disclosure; Page 48; 79pp; English.
XX XX
XX CC The invention discloses analytical reagents (e.g. trifunctional synthetic
XX CC reagents) which can be used for reducing the complexity of peptide
XX CC mixtures. The method labels peptides at a specific amino acid residue and
XX CC then selectively enriches only those peptides containing the labelled
XX CC amino acid. The compound have the formula of immobilisation site-cleavage
XX CC site-link. The immobilisation site is chosen from an epitope tag, a
XX CC linker to a solid surface, a metal chelating site, a magnetic site and a
XX CC specific oligonucleotide sequence, or their combination, the cleavage
XX CC site is chosen from a protease cleavage site, a photocleavable linker, a
XX CC restriction enzyme cleavage site, a chemical cleavage site and a thermal
XX CC cleavage site, or their combination and the link is chosen from an amino
XX CC acid reactive site and a mass variance site, or their combination. The
XX CC compounds are useful for simultaneously identifying and determining the
XX CC levels of expression of cysteine-containing proteins in normal and
XX CC perturbed cells. The advantage is that these reagents allow rapid and
XX CC quantitative analysis of proteins or protein function in mixtures of
XX CC proteins. By preparing the reagent in two forms with detectably different
XX CC masses, accurate relative quantification of peptide amounts using mass
XX CC spectrometry, can be achieved. The sequences given in ABG95935-ABG96244
XX CC are examples of the peptide tags used to isolate cysteine-containing
XX CC proteins, the target sequences tested and the peptides isolated using the
XX CC peptide tags
XX XX
XX SQ Sequence 14 AA;
XX
XX Query Match 36.5%; Score 35; DB 5; Length 14;
XX Best Local Similarity 45.5%; Pred. No. 39;
XX Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 9 LGLSGVQEIKE 19
XX Db ||:|| :||
XX 4 LGITGVQVRE 14
XX
XX RESULT 6
XX RAY07977
XX ID AAY07977 standard; protein; 27 AA.
XX AC AAY07977;
XX XX
XX DT 06-JUL-1999 (first entry)
XX XX
XX DE Human secreted protein fragment #7 encoded from gene 28.
XX KW Human; secreted protein; treatment; prevention; protein therapy; AIDS;
XX KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;
XX KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
XX KW immune system disease; autoimmune disease; hepatic disease; lymphoma;
XX KW renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;
XX KW cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;
XX KW pulmonary disorder; transplant rejection; osteoclast; osteoporosis;
XX KW arthritis; malignancy; digestive; endocrine; infection.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO9918208-A1.
XX XX
XX PD 15-APR-1999.
XX XX

```

05-OCT-2000.
 20-MAR-2000; 2000WO-EP002478.
 26-MAR-1999; 99GB-00007113.
 25-SEP-1999; 99GB-00022858.
 (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 Bruck CEM, Cassart J, Coche T, Vinals Y De BassolsC;
 WPI; 2000-664923/54.
 Novel CASB619 polypeptides useful for diagnosing, and as vaccines for prophylactic and therapeutic treatment of, cancers, particularly ovarian and colon carcinoma, and autoimmune diseases.
 Example 7; Page 62; 68pp; English.
 The present sequence comprises an epitope derived from the human CASB619 protein sequence. This protein is thought to be specifically or over-expressed in tumour cells, and so can be used as a target for antigen-specific immune responses, which can cause destruction of the tumour cell. In addition, the protein and gene can be used in cancer diagnosis, in the treatment of autoimmune diseases and in vaccines against cancer and autoimmune disease. The present sequence can be used as an immunogen

Query Match 33.3%; Score 32; DB 3; Length 10;
 Best Local Similarity 50.0%; Pred. No. 87;
 Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 3 NLADRLMLGLS 12
 Db 1 SLADRLIGVT 10

RESULT 8
 AAB81167
 ID AAB81167 standard; peptide; 20 AA.
 AC AAB81167;
 XX AC
 XX 06-AUG-2003 (revised)
 DT 13-JUL-2001 (first entry)
 XX SVPEP related peptide.
 DE SVPEP related peptide.
 XX Coryneform bacteria; transglutaminase; food processing; svPEP.
 XX Streptomyces mobaraensis.
 OS WO200123591-A1.
 PN 05-APR-2001.
 XX 29-SEP-2000; 2000WO-JP006780.
 PF 30-SEP-1999; 99JP-00280098.
 XX 28-JUN-2000; 2000JP-00194043.
 PR (AJIN) AJINOMOTO CO INC.
 XX Kikuchi Y, Date M, Umezawa Y, Yokoyama K, Matsui H;
 WPI; 2001-266172/27.
 XX Efficient secretory production of foreign proteins e.g. transglutaminase employing transformant coryneform bacterium, simply on industrial scale with direct recovery for use in food processing and pharmaceutical industry.

01-OCT-1998; 98WO-US020775.
 02-OCT-1997; 97US-0060833P.
 02-OCT-1997; 97US-0060836P.
 02-OCT-1997; 97US-0060837P.
 02-OCT-1997; 97US-0060838P.
 02-OCT-1997; 97US-0060839P.
 02-OCT-1997; 97US-0060843P.
 02-OCT-1997; 97US-0060862P.
 02-OCT-1997; 97US-0060874P.
 02-OCT-1997; 97US-0060880P.
 02-OCT-1997; 97US-0060884P.
 (HUMA-) HUMAN GENOME SCI INC.
 Duan DR, Florence KA, Rosen CA, Ruben SM, Greene JM, Young P;
 Ferrite AM, Yu G, Janat F, Ni J, Carter KC, Endress GA, Feng P;
 Lafleur DW, Shi Y;
 WPI; 1999-264022/22.
 New isolated human genes and the secreted polypeptides they encode.
 Disclosure; Page 338; 369pp; English.
 This invention describes novel isolated human genes and the secreted proteins they encode. The products of the invention are useful for preventing, treating or ameliorating medical conditions, e.g. by protein or gene therapy. Also pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new polynucleotides. Specific uses are described for each of the 101 polynucleotides, based on which tissues they are most highly expressed in, and include developing products for the diagnosis or treatment of cancer, tumours, neurodegenerative disorders, developmental abnormalities and fetal deficiencies, blood disorders, leukemias, diseases of the immune system, autoimmune diseases, hepatic and renal disease, lymphomas, inflammation, allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate disease, skeletal or cardiac muscle disorders, pulmonary disorders, transplant rejection, disorders involving osteoclasts such as osteoporosis, arthritis or malignancies, digestive/endocrine disorders, infections and AIDS. The human secreted proteins of the invention are represented in AAY07852-Y07993 and the encoding nucleic acids are represented in AAX37451-X37552

Query Match 34.4%; Score 33; DB 2; Length 27;
 Best Local Similarity 56.2%; Pred. No. 1.9e+02;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 INLADRLMLGLSGVQEI 17
 Db 2 IHLEKRSGLSETQII 17

RESULT 7
 AAB27118
 ID AAB27118 standard; protein; 10 AA.
 AC AAB27118;
 XX 12-FEB-2001 (first entry)
 DT Human CASB619 protein epitope SEQ ID NO: 44.
 DE Human; CASB619; cancer; autoimmune disease; immunogen; vaccine; epitope.
 XX Homo sapiens.
 OS WO200058460-A2.
 PN

PS Example 9; Page 50-57; 151pp; Japanese.

CC This invention relates to a process for the production of a foreign

CC secretory protein through the construction of a recombinant coryneform

CC bacterium. The coryneform bacterium is transformed with an expression

CC construct in which DNA encoding a target foreign protein pro-structure is

CC ligated to the downstream region of DNA encoding the signal peptide

CC domain of a coryneform bacterial protein. Following transformation with

CC the vector, the bacterium is cultured, and the pro-peptide cleaved from

CC the expressed protein. Transglutaminases produced using this process are

CC useful in the food processing and pharmaceutical industries. The present

CC sequence represents a svPEP related peptide used in an example

CC illustrating the method of the invention. (Updated on 06-AUG-2003 to

CC correct OS field.)

XX

XX Sequence 20 AA;

Query Match 33.3%; Score 32; DB 4; Length 20;

Best Local Similarity 27.8%; Pred. No. 2e+02;

Matches 5; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 3 NLADRLGLSGVQEIKEQ 20

Db 3 DIKDRILKIPGMKFVEEK 20

RESULT 9

ABU07394

ID ABU07394 standard; peptide; 20 AA.

XX

AC ABU07394;

XX

XX

DT 23-OCT-2003 (revised)

DT 28-JAN-2003 (first entry)

XX

DE Foreign protein #5.

XX

KW Coryneform bacterium; signal peptide domain; food processing; medicine;

KW cosmetic; transglutaminase; human epithelial growth factor.

XX

OS Streptomyces mobaraensis.

XX

PN WO200281694-A1.

XX

PD 17-OCT-2002.

XX

PF 27-MAR-2002; 2002WO-JP002978.

XX

PR 30-MAR-2001; 2001JP-00098808.

XX

PA (AJIN) AJINOMOTO CO INC.

XX

PI Kikuchi Y, Date M, Umezawa Y, Yokoyama K, Heima H, Matsui H;

XX

DR WPI; 2003-058550/05.

XX

PT Secretion production of foreign proteins by culturing transformant

PT coryneform bacteria, applicable in producing e.g. industrially-useful

PT transglutaminase and human epithelial growth factor.

XX

PS Example 7; Page 39; 117pp; Japanese.

XX

CC The invention relates to a method for producing a foreign protein by

CC culturing a mutant strain of a coryneform bacterium that contains an

CC expression gene construct to secrete the foreign protein, followed by

CC recovery of the produced foreign protein. The gene construct is obtained

CC by ligation of a nucleic acid encoding a signal peptide domain

CC originating from a coryneform bacterium, to downstream of a functioning

CC promoter sequence in the coryneform bacterium and also by ligation of a

CC nucleic acid encoding a foreign protein, to downstream of a nucleic acid

CC sequence encoding the signal peptide. The method is useful for the

CC production of a foreign protein which is applicable in producing e.g.

CC industrially-useful transglutaminase and human epithelial growth factor

CC

CC for use in medicine, cosmetics and food processing. This sequence

CC represents a foreign protein of the invention. (Updated on 23-OCT-2003 to

CC standardise OS field)

XX

SQ Sequence 20 AA;

Query Match 33.3%; Score 32; DB 6; Length 20;

Best Local Similarity 27.8%; Pred. No. 2e+02;

Matches 5; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 3 NLADRLGLSGVQEIKEQ 20

Db 3 DIKDRILKIPGMKFVEEK 20

RESULT 10

ADD13104

ID ADD13104 standard; peptide; 23 AA.

XX

AC ADD13104;

XX

DT 01-JAN-2004 (first entry)

XX

DE Human LDH-C protein SEQ ID 17.

XX

KW tumour-associated antigen; TAG; tumour-inhibiting activity; HLA;

KW human leukocyte antigen; metastatic cancer; TPTE; cytostatic; vaccine;

KW cancer; lung; breast; colon; kidney; cervix; melanoma; human; LDH-C.

XX

OS Homo sapiens.

XX

PN WO2003076631-A2.

XX

PD 18-SEP-2003.

XX

PF 12-MAR-2003; 2003WO-EP002556.

XX

PR 13-MAR-2002; 2002DE-01011088.

XX

PA (SAHI/) SAHIN U.

PA (TUER/) TUERECI O.

PA (KOSL/) KOSLOWSKI M.

XX

PI Sahin U, Tuereci O, Koslowski M;

XX

DR WPI; 2003-756825/71.

XX

PT Compositions containing inhibitors of specific tumor antigens, useful for

PT diagnosis, treatment and monitoring of cancers, and related nucleic

PT acids, proteins and antibodies.

XX

PS Claim 32; SEQ ID NO 17; 173pp; German.

XX

CC This invention describes a novel pharmaceutical composition containing an

CC agent that (a) inhibits expression or activity of a tumour-associated

CC antigen (TAG), (b) has selective tumour-inhibiting activity against cells

CC with (abnormal) expression of TAG, or (c) increases the amount of complex

CC formed between an HLA (human leukocyte antigen) molecule and TAG, or its

CC fragments. The products of the invention are used for (1) diagnosing

CC diseases associated with (abnormal) expression of TAG, (2) determining

CC regression, progression or outbreak of diseases associated with

CC (abnormal) expression of TAG, (3) treating diseases associated with

CC (abnormal) expression of TAG and (4) treatment, diagnosis and monitoring

CC of metastatic cancer associated with (abnormal) expression of TPTE. The

CC products of the invention have cytostatic activity and can be used as

CC vaccines or for antisense inhibition. The compositions are used for

CC treatment, diagnosis and monitoring of cancer (of lung, breast, colon

CC (including metastases), kidney or cervix, also melanoma). This sequence

CC represents a human LDH-C protein fragment.

XX

SQ Sequence 23 AA;

Query Match 33.3%; Score 32; DB 7; Length 23;

Best Local Similarity 36.8%; Pred. No. 2.3e+02; Mismatches 5; Indels 7; Gaps 0;

Matches 7; Conservative

QY 2 INLADRLGLSGVQEIKEQ 20

DB 2 VGLLENVILVGLYGIKEE 20

ABG62321

ABG62321 standard; peptide; 25 AA.

AC ABG62321;

DT 21-AUG-2002 (first entry)

DE Eubacterial DNA polymerase IV QLSLF motif containing peptide #51.

XX DNA polymerase III; beta subunit; eubacteria; antibacterial;

XX eubacterial infection.

XX Mycobacterium avium.

XX WO200238596-A1.

XX 16-MAY-2002.

XX 08-NOV-2001; 2001WO-AU001436.

XX 08-NOV-2000; 2000AU-00001320.

XX 06-FEB-2001; 2001AU-00002919.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX Dalrymple BP, Konguwan K, Wijffels GL, Jennings PA, Kemp GW;

XX WPT; 2002-471546/50.

XX New molecule having surface analogous to surface of domain of eubacterial

XX beta protein contacted by proteins that interact with beta protein,

XX useful to identify inhibitors of beta protein-ligand interaction.

XX Example 1; Page 27; 326pp; English.

XX The invention relates to a molecule (I) comprising a surface (S)

XX analogous to the surface of the domain of eubacterial beta protein

XX contacted by proteins that interact with beta protein, where the surface

XX is defined by the residues X(170), X(172), X(175), X(177), X(241),

XX X(242), X(247), X(346), X(360), and X(362), where the superscript numbers

XX designate the position of residues in Escherichia coli beta protein, or

XX the equivalent residues in homologues from other species of eubacteria,

XX and where: X(170) = Val, Ile, Ala, Thr, Ser or Glu; X(172) = Thr, Ser or

XX Ile; X(175) = His, Tyr, Phe, Lys, Ile, Glu or Arg; X(177) = Leu, Met,

XX Ile, Phe, Val or Ala; X(241) = Phe, Tyr or Leu; X(242) = Pro, Leu or Ile;

XX X(247) = Val, Ile, Ala, Phe, Leu or Met; X(346) = Ser, Pro, Ala, Tyr or

XX Lys; X(360) = Ile, Leu or Val; and X(362) = Met, Leu, Val, Ser, Thr or

XX Arg. Also included are methods of identifying a modulator of the

XX interaction between a eubacterial beta protein and proteins that interact

XX with them, reducing (M4) the effect of eubacterial infestation of a

XX biological system, involves delivering to a system infested with a

XX eubacterial species, a modulator of the interaction between eubacterial

XX beta protein and proteins that interact with the beta protein; and (4) a

XX template (II) for the design of a compound that binds to at least part of

XX (S) of beta protein as defined above comprises a (P) such as X¹IX², where:

XX X¹ = any amino acid residue; X² = Leu, Ile, or Phe; X³ = Ala, Gly, Thr, Asn, Asp,

XX Cys, Phe, Tyr, Trp, Pro, Asp, Ala or Gly; X⁴ = Leu, Ile, Val, Cys,

XX Ser, or Pro; X⁵ = Ala or Gly; X⁶ = Leu; and X⁷ = Leu, Ile, Val, Cys,

XX Phe, Tyr, Trp or Pro. The method are useful for identifying a modulator

XX of the interaction between a eubacterial beta protein and proteins that

XX interact with the beta protein. (M4) is useful for reducing the effect of

XX eubacterial infestation of a biological system. The compounds identified

XX using above mentioned methods are useful as antibacterial agent for

CC treatment or prevention of disease in humans, animals and plants. The

CC present sequence is a eubacterial peptide from a DNA binding protein or

CC polymerase which contains a DNA polymerase III beta subunit binding site

XX SQ Sequence 25 AA;

Query Match 33.3%; Score 32; DB 5; Length 25;

Best Local Similarity 33.3%; Pred. No. 2.5e+02;

Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 9 LGLSGVQEIKEQ 20

DB 1 VGFSGLSVQEQ 12

RESULT 12

AAW77325

ID AAW77325 standard; peptide; 28 AA.

XX AC AAW77325;

XX 14-DEC-1998 (first entry)

XX Human ReproSA-1 peptide 2.

XX Human; ReproSA-1; contraceptive; anti-sperm; fertility;

XX in vitro fertilisation.

XX Homo sapiens.

XX WO9836073-A1.

XX 20-AUG-1998.

XX 17-FEB-1998; 98WO-US003513.

XX 18-FEB-1997; 97US-0039574P.

XX (REPR-) REPROGEN INC.

XX French CK, Neilson LI, Schneider PA, Yamamoto KK;

XX WPI; 1998-480771/41.

XX Sperm-specific polypeptide ReproSA-1 and related nucleic acid -

XX transformed cells and antibodies, useful in contraceptive vaccines or

XX compositions, and for diagnosis and treatment of infertility.

XX Claim 7; Page 47; 62pp; English.

XX The ReproSA-1 polypeptide, it's fragments, or nucleic acid encoding them

XX are used in contraceptive vaccines to generate an anti-sperm response.

XX Antibodies can also be used as topical contraceptive (spermicide), also

XX antisense sequences, ribozymes and triplex-forming molecules. Detection,

XX by reaction with immobilised peptides, of anti-(I) antibodies (in serum,

XX semen, saliva, cervical or vaginal mucosa) is used to diagnose

XX immunological infertility and to monitor the effect of vaccination.

XX nucleotide fragments are used to diagnose infertility-associated

XX mutations in the gene encoding ReproSA, and for isolation of related

XX sequences. Peptides and anti-idiotypic antibodies against the antibodies

XX are used to inactivate the antibodies i.e. to increase fertility and may

XX be added to sperm samples before in vitro fertilisation

XX SQ Sequence 28 AA;

Query Match 33.3%; Score 32; DB 2; Length 28;

Best Local Similarity 63.6%; Pred. No. 2.9e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 10 GLSGVQEIKEQ 20

DB 16 GLKKVQEIKEQ 26

```

RESULT 13
AAB26182
ID AAB26182 standard; protein; 9 AA.
XX
AC AAB26182;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human CASB619 protein epitope SEQ ID NO: 5.
XX
KW Human; CASB619; cancer; autoimmune disease; immunogen; vaccine; epitope.
XX
OS Homo sapiens.
XX
PN WO200058460-A2.
XX
PD 05-OCT-2000.
XX
PF 20-MAR-2000; 2000WO-EP002478.
XX
PR 26-MAR-1999; 99GB-00007113.
XX
PR 25-SEP-1999; 99GB-00022858.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Bruck CEM, Cassart J, Coche T, Vinals Y De BassolsC;
XX
DR WPI; 2000-664923/64.
XX
PT Novel CASB619 polypeptides useful for diagnosing, and as vaccines for
PT prophylactic and therapeutic treatment of, cancers, particularly ovarian
PT and colon carcinoma, and autoimmune diseases.
XX
PS Example 7; Page 57; 68pp; English.
XX
CC The present sequence comprises an epitope derived from the human CASB619
CC protein sequence. This protein is thought to be specifically or over-
CC expressed in tumour cells, and so can be used as a target for antigen-
CC specific immune responses which can cause destruction of the tumour cell.
CC In addition, the protein and gene can be used in cancer diagnosis, in the
CC treatment of autoimmune diseases and in vaccines against cancer and
CC autoimmune disease. The present sequence can be used as an immunogen
XX
SQ Sequence 9 AA;
XX
Query Match 32.3%; Score 31; DB 3; Length 9;
Best Local Similarity 55.6%; Pred. No. 1.4e+06;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 3 NLADRLGL 11
Db :|||||:
1 SLADRLGLV 9
XX
RESULT 14
AAO18055
ID AAO18055 standard; peptide; 14 AA.
XX
AC AAO18055;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human neuropeptide Y signal peptide fragment #1.
XX
KW Human; neuropeptide Y; NPY; chromosome 7q15.1; SNP; obesity; BMI;
KW body mass index; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200236825-A1.
XX
PD 10-MAY-2002.
XX
Query Match 32.3%; Score 31; DB 3; Length 9;
Best Local Similarity 55.6%; Pred. No. 1.4e+06;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 3 NLADRLGL 11
Db :|||||:
1 SLADRLGLV 9
XX
RESULT 15
AAO18057
ID AAO18057 standard; peptide; 14 AA.
XX
AC AAO18057;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human neuropeptide Y signal peptide fragment #2.
XX
KW Human; neuropeptide Y; NPY; chromosome 7q15.1; SNP; obesity; BMI;
KW body mass index; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200236825-A1.
XX
PD 10-MAY-2002.
XX
PF 02-NOV-2001; 2001WO-SE002431.
XX
PR 03-NOV-2000; 2000SE-00004035.
XX
PA (PHAA ) PHARMACIA AB.
XX
PI Wahlestedt C, Ding B;
XX
DR WPI; 2002-500129/53.
XX
PT Diagnosing single nucleotide polymorphism(s) in the human neuropeptide Y
PT (NPY) gene, useful for diagnosing a predisposition to e.g. obesity, by
PT determining the nucleic acid sequence at one or more positions of the NPY
PT gene in the human.
XX
PS Example 3; Page 10; 37pp; English.
XX
CC The present invention relates to a method of diagnosing a predisposition
CC to obesity in humans by identifying a single nucleotide polymorphism
CC (SNP) at position 1128 (position 106 of this sequence) of the
CC neuropeptide Y gene. The gene is found at chromosome 7q15.1. The present
CC sequence is a fragment of the NPY signal peptide
XX
SQ Sequence 14 AA;
XX
Query Match 32.3%; Score 31; DB 5; Length 14;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 4 LADRLGLSGV 14
Db :|||||:
2 LGNKLGLSL 12
XX
RESULT 16
AAO18057
ID AAO18057 standard; peptide; 14 AA.
XX
AC AAO18057;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human neuropeptide Y signal peptide fragment #2.
XX
KW Human; neuropeptide Y; NPY; chromosome 7q15.1; SNP; obesity; BMI;
KW body mass index; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200236825-A1.
XX
PD 10-MAY-2002.
XX
PF 02-NOV-2001; 2001WO-SE002431.
XX
PR 03-NOV-2000; 2000SE-00004035.
XX
PA (PHAA ) PHARMACIA AB.
XX
PI Wahlestedt C, Ding B;
XX
DR WPI; 2002-500129/53.
XX
PT Diagnosing single nucleotide polymorphism(s) in the human neuropeptide Y
PT (NPY) gene, useful for diagnosing a predisposition to e.g. obesity, by
PT determining the nucleic acid sequence at one or more positions of the NPY
PT gene in the human.
XX
PS Example 3; Page 10; 37pp; English.
XX
CC The present invention relates to a method of diagnosing a predisposition
CC to obesity in humans by identifying a single nucleotide polymorphism

```

CC (SNP) at position 1128 (position 106 of this sequence) of the
CC neurotrophin-4 gene. The gene is found at chromosome 7q15.1. The present
CC sequence is a fragment of the NPY signal peptide
XX
SQ Sequence 14 AA;
Query Match 32.3%; Score 31; DB 5; Length 14;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 4 LADRLGLSGV 14
| : : : : :
Db 2 LGNRLGLSGL 12
Search completed: July 15, 2004, 14:46:51
Job time : 43.1333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:40:36 ; Search time 8.5333 Seconds
(without alignments)
225.449 Million cell updates/sec

Title: US-09-171-432A-46
Perfect score: 96
Sequence: 1 KINLADRLGLSGVQEIKEQ 20
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 6282

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	31	32.3	28	2	S55729	otridine-5'-monop
2	29	30.2	25	2	PC1314	large granule L1 c
3	27	28.1	12	2	S65626	phosphorylase b -
4	26	27.1	19	2	A61243	dimethylaniline mo
5	26	27.1	20	2	S35970	ribosomal protein
6	26	27.1	21	2	S35978	ribosomal protein
7	26	27.1	23	2	S35975	ribosomal protein
8	26	27.1	23	2	S35976	ribosomal protein
9	26	27.1	24	2	S20273	photosystem II 10K
10	25.5	26.6	26	2	TL5472	hypothetical prote
11	25	26.0	20	2	C49753	hypothetical prote
12	25	26.0	25	2	D47689	flagellar core pro
13	25	26.0	25	2	S68005	hucolin, 35K chain
14	25	26.0	27	2	E44621	homeotic protein H
15	25	26.0	30	2	B70165	hypothetical prote
16	25	26.0	30	2	H84074	hypothetical prote
17	24	25.0	30	2	JP0055	ribosomal protein
18	24	25.0	32	2	S70780	aspartate transami
19	24	25.0	23	2	S43469	heterodisulfide re
20	24	25.0	28	2	S47624	D-aspartate oxidas
21	24	25.0	28	2	A44877	cell surface prote
22	23	24.0	14	2	G61308	hemocyanin chain 3
23	23	24.0	15	2	PA0075	fructose-bisphosph
24	23	24.0	15	2	PA0102	fructose-bisphosph
25	23	24.0	17	2	S51736	T-cell receptor be
26	23	24.0	18	2	I49408	cytochrome-c oxida
27	23	24.0	19	2	SI0747	dimethylaniline mo
28	23	24.0	22	2	A35418	brain natriuretic
29	23	24.0	22	2	A45913	plantaricin A - La

ALIGNMENTS

RESULT 1

S55729
otridine-5'-monophosphate decarboxylase - Aspergillus awamori (fragments)
C:Species: Aspergillus awamori
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S55729
R:Gouka, R.J.; Hessing, J.G.M.; Stam, H.; Musters, W.; van den Hondel, C.A.M.J.J.
Curr. Genet. 27, 536-540, 1995
A:Title: A novel strategy for the isolation of defined pyrG mutants and the development
A:Reference number: S55729; PMID:96031709; PMID:7553938
A:Accession: S55729
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-18;19-28 <GOU>
C:Superfamily: orotidine-5'-phosphate decarboxylase; orotidine-5'-phosphate decarboxylase

Query Match 32.3%; Score 31; DB 2; Length 28;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 INLADRLGL 10
Db 14 LIDLADRVNG 22

RESULT 2

PC1314
large granule L1 chain - horseshoe crab (Tachypleus tridentatus) (fragment)
C:Species: Tachypleus tridentatus
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
C:Accession: PC1314
R:Shigenaga, T.; Takayevnoki, Y.; Kawasaki, S.; Seki, N.; Muta, T.; Toh, Y.; Ito, A.; Iwar
J. Biochem. 114, 307-316, 1993
A:Title: Separation of large and small granules from horseshoe crab (Tachypleus tridentat
A:Reference number: PC1309; PMID:94110249; PMID:8282718
A:Accession: PC1314
A:Molecule type: protein
A:Residues: 1-25 <SHI>
C:Comment: This protein participates in immobilization of invading microbes.

Query Match 30.2%; Score 29; DB 2; Length 25;
Best Local Similarity 46.7%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 NLADRLGLSGVQEI 17
Db 5 NKMEMLXDSGDQEV 19

RESULT 3

S65626
phosphorylase b - rabbit (fragments)

cytochrome-c oxida
chaperonin, 10K -
14K aggregative ad
NADP phosphatase I
glyceraldhyde-3-p
serine proteinase
hook protein - Cam
homeotic protein H
brain natriuretic
trypsin-like serin
outer membrane pro
NADH2 dehydrogenas
sialokin I - yel
Ig heavy chain DJ
GLYMA1 - soybean (

30 23 24.0 23 2 S43289
31 23 24.0 23 2 B39313
32 23 24.0 23 2 A49221
33 23 24.0 23 2 A59480
34 23 24.0 24 2 PN0163
35 23 24.0 25 2 A60704
36 23 24.0 26 2 A44910
37 23 24.0 27 2 JCI081
38 23 24.0 27 2 B44636
39 23 24.0 30 2 A32946
40 22.5 23.4 20 2 S06802
41 22 22.9 7 2 T09512
42 22 22.9 10 2 A49581
43 22 22.9 14 2 PHI321
44 22 22.9 15 2 A56970
45 22 22.9 16 2 S42237

C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C:Accession: S65626
R:Luo, S.; Martin, B.L.; Senshu, T.; Graves, D.J.
Arch. Biochem. Biophys. 318, 362-369, 1995
A:Title: Enzymatic deamination of glycogen phosphorylase and a peptide of the phosphorylase
A:Reference number: S65626; MUID:95251385; PMID:7733664
A:Accession: S65626
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-3;4-12 <LUO>

Query Match 27.1%; Score 27; DB 2; Length 12;
Best Local Similarity 44.4%; Pred. No. 3.2e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 LGLSGVQVEI 17
DB 3 VGLAGVENV 11

RESULT 4
A61243
dimethylamine monooxygenase (N-oxide-forming) (EC 1.14.13.8), renal - mouse (fragment)
N:Alternate names: dimethylamine oxidase; flavin-containing monooxygenase
C:Species: Mus musculus (house mouse)
C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 27-Jun-1994
C:Accession: A61243
R:Venkatesh, K.; Levi, P.E.; Hodgson, E.
Biochem. Pharmacol. 42, 1411-1420, 1991
A:Title: The flavin-containing monooxygenase of mouse kidney. A comparison with the live
A:Reference number: A61243; MUID:92029024; PMID:1930264
A:Accession: A61243
A:Molecule type: protein
A:Residues: 1-19 <VEN>
C:Keywords: FAD; flavoprotein; microsome; monooxygenase; NADP; oxidoreductase

Query Match 27.1%; Score 26; DB 2; Length 19;
Best Local Similarity 55.6%; Pred. No. 7.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 GLSGVQVEIK 18
DB 10 GVSLASIK 18

RESULT 5
S35970
ribosomal protein L10 - Citrobacter freundii (fragment)
C:Species: Citrobacter freundii
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Oct-1997
C:Accession: S35970
R:Zhyvoloup, A.
submitted to the EMBL Data Library, August 1993
A:Reference number: S35970
A:Accession: S35970
A:Molecule type: DNA
A:Residues: 1-20 <ZHY>
A:Cross-references: EMBL:X74448
C:Genetics:
A:Gene: rplJ
C:Superfamily: Escherichia coli ribosomal protein L10
C:Keywords: protein biosynthesis; ribosome

Query Match 27.1%; Score 26; DB 2; Length 20;
Best Local Similarity 31.2%; Pred. No. 8.1e+02;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 INLADRLCLSGVQEI 17
DB 3 LNLQDKQAIIVAEVSEV 18

C:Species: Proteus vulgaris (fragment)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 13-Aug-1999
C:Accession: S35978
R:Zhyvoloup, A.
submitted to the EMBL Data Library, August 1993
A:Reference number: S35978
A:Accession: S35978
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-21 <ZHY>
A:Cross-references: EMBL:X74446; NID:g396471; PIDN:CRA52457.1; PID:g396472
C:Genetics:
A:Gene: rplJ
C:Superfamily: Escherichia coli ribosomal protein L10
C:Keywords: protein biosynthesis; ribosome

Query Match 27.1%; Score 26; DB 2; Length 21;
Best Local Similarity 31.2%; Pred. No. 8.6e+02;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 INLADRLCLSGVQEI 17
DB 3 LNLQDKQAIIVAEVSEV 18

RESULT 7
S35975
ribosomal protein L10 - Enterobacter cloacae (fragment)
C:Species: Enterobacter cloacae
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 29-Aug-1997
C:Accession: S35975
R:Zhyvoloup, A.
submitted to the EMBL Data Library, August 1993
A:Reference number: S35975
A:Accession: S35975
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-23 <ZHY>
A:Cross-references: EMBL:X74444
C:Genetics:
A:Gene: rplJ
C:Superfamily: Escherichia coli ribosomal protein L10
C:Keywords: protein biosynthesis; ribosome

Query Match 27.1%; Score 26; DB 2; Length 23;
Best Local Similarity 31.2%; Pred. No. 9.5e+02;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 INLADRLCLSGVQEI 17
DB 3 LNLQDKQAIIVAEVSEV 18

RESULT 8
S35976
ribosomal protein L10 - Klebsiella pneumoniae (fragment)
C:Species: Klebsiella pneumoniae
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 29-Aug-1997
C:Accession: S35976
R:Zhyvoloup, A.
submitted to the EMBL Data Library, August 1993
A:Reference number: S35976
A:Accession: S35976
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-23 <ZHY>
A:Cross-references: EMBL:X74445
C:Genetics:
A:Gene: rplJ
C:Superfamily: Escherichia coli ribosomal protein L10

C;Keywords: protein biosynthesis; ribosome

Query Match 27.1%; Score 26; DB 2; Length 23;
Best Local Similarity 31.2%; Pred. No. 9.5e+02;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 INLADRMGLSGVQRI 17

Db 3 LNLQDKQAIVAEVSEV 18

RESULT 9

S02073

photosystem II 10K protein precursor - wheat (fragment)

C;Species: Triticum aestivum (common wheat)

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 23-Jul-1999

C;Accession: S02073

R;Webber, A.N.; Packman, L.C.; Gray, J.C.

FEBS Lett. 242, 435-438, 1989

A;Title: A 10 kDa polypeptide associated with the oxygen-evolving complex of photosystem II

A;Reference number: S02073; PMID:89121115; PMID:2644135

A;Accession: S02073

A;Molecule type: protein

A;Residues: 1-24 <WEB>

C;Superfamily: photosystem II 10K protein

C;Keywords: chloroplast; photosynthesis; photosystem II; thylakoid

Query Match 27.1%; Score 26; DB 2; Length 24;

Best Local Similarity 62.5%; Pred. No. 9.9e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 DRMLGLSG 13

Db 9 DKPLGLGG 16

RESULT 10

TI15472

hypothetical protein C09B9.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: TI15472

R;Bradshaw, H.

submitted to the EMBL Data Library, February 1996

A;Description: The sequence of C. elegans cosmid C09B9.

A;Reference number: Z18355

A;Accession: TI15472

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-26 <BRA>

A;Cross-references: EMBL:U50069; NID:gl702977; PID:gl208866; PIDN:AAB37556.1; GSPDB:GN00

A;Experimental source: strain Bristol N2; clone C09B9

C;Genetics:

A;Gene: CESP:C09B9.5

A;Map position: 4

A;Introns: 12/1

Query Match 26.6%; Score 25.5; DB 2; Length 26;

Best Local Similarity 44.4%; Pred. No. 1.3e+03;

Matches 8; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 2 INLADRMGLSGVQRI 19

Db 3 VNWRD---GGGCVGEFKE 17

RESULT 11

C49753

hypothetical protein (proB 5' region) - Serratia marcescens (fragment)

C;Species: Serratia marcescens

C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 08-Oct-1999

C;Accession: C49753; S11643

R;Omori, K.; Suzuki, S.I.; Imai, Y.; Komatsubara, S.

J. Gen. Microbiol. 137, 509-517, 1991

A;Title: Analysis of the Serratia marcescens proBA operon and feedback control of proline

A;Reference number: A49753; PMID:91237315; PMID:1851803

A;Accession: C49753

A;Molecule type: DNA

A;Residues: 1-20 <OMO>

A;Cross-references: GB:X53086; NID:g47251; PIDN:CAA37253.1; PID:g47252

Query Match 26.0%; Score 25; DB 2; Length 20;

Best Local Similarity 41.7%; Pred. No. 1.2e+03;

Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 9 LGLSGVQRIKEQ 20

Db 3 LGLEPADDFKEK 14

RESULT 12

D47689

flagellar core protein, 32K - Treponema hyodysenteriae (fragment)

C;Species: Treponema hyodysenteriae

C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 24-Feb-1995

C;Accession: D47689

R;Koopman, M.B.; Baats, E.; van Vorstenbosch, C.J.; van der Zeijst, B.A.; Kusters, J.G.

J. Gen. Microbiol. 138, 2697-2706, 1992

A;Title: The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are composed of

A;Reference number: A47689; PMID:93139764; PMID:1487733

A;Contents: C5, Treponema

A;Accession: D47689

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-25 <KOO>

A;Note: sequence extracted from NCBI backbone (NCBIF:123403)

Query Match 26.0%; Score 25; DB 2; Length 25;

Best Local Similarity 55.6%; Pred. No. 1.5e+03;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 ADRMLGLSG 13

Db 12 ANRQLNLGTG 20

RESULT 13

S68005

hucolin, 35K chain - human (fragment)

N;Alternate names: 36K HLA-cross-reactive plasma protein

C;Species: Homo sapiens (man)

C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 18-Feb-2000

C;Accession: S68005; A38648

R;Edgar, P.F.

FEBS Lett. 375, 159-161, 1995

A;Title: Hucolin, a new corticosteroid-binding protein from human plasma with structural

A;Reference number: S68004; PMID:96087107; PMID:7498469

A;Accession: S68005

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-25 <EDG>

R;Hada, J.A.; She, J.X.; Kao, K.J.

J. Biol. Chem. 266, 3695-3701, 1991

A;Title: Biochemical characterization of 39-kDa class I histocompatibility antigen in pl

A;Reference number: A38648; PMID:91139661; PMID:1995624

A;Accession: A38648

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-6 <HAG>

Query Match 26.0%; Score 25; DB 2; Length 25;

Best Local Similarity 36.4%; Pred. No. 1.5e+03;

Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 RMLGLSGVQRI 17

Db 17 RMLGLSGVQRI 17

Db 11 KMVLEGSDDL 21

RESULT 14

E44621
homeotic protein Hox 10 (clone 10w) - sea lamprey (fragment)
C; Species: Petromyzon marinus (sea lamprey)
C; Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 15-Oct-1999
C; Accession: E44621
R; Bendleton, J.W.; Nagai, B.K.; Murtha, M.T.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 90, 6300-6304, 1993
A; Title: Expansion of the Hox gene family and the evolution of chordates.
A; Reference number: A44616; MUID:93317663; PMID:8101001
A; Accession: E44621
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-27 <PEN>
A; Cross-references: GB:L14895; NID:G290813; PID:AAA02530.1; PID:G290814
C; Superfamily: unassigned homeobox proteins; homeobox homology
C; Keywords: DNA binding; embryo; homeobox; nucleus; transcription regulation

Query Match 26.0%; Score 25; DB 2; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 INLADRL 9

Db 18 VNLTRQV 25

RESULT 15

B70165
hypothetical protein BB0523 - Lyme disease spirochete
C; Species: Borrelia burgdorferi (Lyme disease spirochete)
C; Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C; Accession: B70165
R; Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A; Authors: Smith, H.O.; Venter, J.C.
A; Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A; Reference number: A70100; MUID:98065943; PMID:9403685
A; Accession: B70165
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-30 <KLE>
A; Cross-references: GB:AF001154; GB:AE000783; NID:G2688431; PID:AA06894.1; PID:G268844
A; Experimental source: strain B31

Query Match 26.0%; Score 25; DB 2; Length 30;
Best Local Similarity 25.0%; Pred. No. 1.9e+03;
Matches 3; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 9 LGLSGVQRIKEQ 20

Db 1 MGFNGILAKQR 12

Search completed: July 15, 2004, 14:52:14
Job time : 9.53333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:01 ; Search time 5.3333 Seconds
(without alignments)
195.263 Million cell updates/sec

Title: US-09-171-432A-46
Perfect score: 96
Sequence: 1 KINLADRLGLSGVQEIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 2073

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	26	27.1	19	1 RL10_CITFR	P43448 citrobacter
2	26	27.1	20	1 RL10_PROVU	P51411 proteus vul
3	26	27.1	22	1 RL10_KLEPN	P41190 klebsiella
4	26	27.1	23	1 RL10_ENTCL	Q47608 enterobacte
5	26	27.1	24	1 PSBR_WHEAT	P12358 triticum ae
6	25	26.0	13	1 IDRC_PIG	P20304 sus scrofa
7	25	26.0	20	1 YPRC_SERMA	P22581 serratia ma
8	25	26.0	25	1 FLB3_TREHY	P80161 treponema h
9	25	26.0	30	1 Y523_BORBU	O51473 borrelia bu
10	24	25.0	29	1 GALA_ONCMY	P47213 oncorhynch
11	24	24.0	12	1 CD11_LITXA	P56245 litoria xan
12	23	24.0	19	1 FM01_CAVPO	P49328 cavia porce
13	23	24.0	19	1 UP27_UPEMJ	P82039 uperoleia m
14	23	24.0	23	1 NDP2_ARTSK	P83576 arthrobacte
15	23	24.0	25	1 CR22_LITGI	P56234 litoria gil
16	23	24.0	25	1 CR25_LITGI	P56237 litoria gil
17	23	24.0	27	1 MDH_RHURU	P80459 rhodospirill
18	22	22.9	10	1 TRS1_AEDAE	P42634 aedes aegypt
19	22	22.9	13	1 CD71_LITEW	P82051 litoria ewi
20	22	22.9	16	1 AU26_LITRA	P82393 litoria ran
21	22	22.9	23	1 AU41_LITAU	P82397 litoria aur
22	22	22.9	24	1 AMAA_BACTR	P37356 bacillus th
23	22	22.9	27	1 MD02_PSEMR	P83447 pseudonanas
24	21	21.9	12	1 FRE1_LITIN	P82021 litoria inf
25	21	21.9	18	1 RIP_SIRGR	P83323 siraitia gr
26	21	21.9	19	1 FLA3_SPIAU	P21986 spirochaeta
27	21	21.9	19	1 UP22_UPEIN	P82028 uperoleia i
28	21	21.9	20	1 CA01_RAT	P19633 rattus norv
29	21	21.9	22	1 ANFC_CHICK	P21805 gallus gall
30	21	21.9	22	1 AOFM_MOUSE	Q64133 mus musculus
31	21	21.9	22	1 MOTI_CANFA	P19863 canis famli
32	21	21.9	23	1 NIFD_ANASL	P33177 anabaena sp
33	21	21.9	25	1 CR24_LITCE	P56236 litoria cae

34	21	21.9	27	1 FLA4_SPIAU	P21987 spirochaeta
35	20.5	21.4	25	1 DNAJ_MYCCA	P71500 mycoplasma
36	20.5	21.4	28	1 PETL_CYAPA	P48102 cyanophora
37	20	20.8	10	1 TKL4_LOCFI	P30250 locusta mig
38	20	20.8	11	1 RE41_LITRU	P82074 litoria rub
39	20	20.8	11	1 TKNA_RANCA	P22688 rana catesb
40	20	20.8	12	1 CD14_LITXA	P56246 litoria xan
41	20	20.8	20	1 PSAF_PEA	P20119 pisum sativ
42	20	20.8	21	1 CHIC_PEA	P31233 pisum sativ
43	20	20.8	21	1 SYA_RAT	P50475 rattus norv
44	20	20.8	22	1 DNI4_BPD10	Q38199 bacterioph
45	20	20.8	22	1 MOTI_CHICK	Q9prp6 gallus gall

ALIGNMENTS

RESULT 1
RL10_CITFR STANDARD; PRT; 19 AA.
AC P43448;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10 (Fragment).
GN RPLJ.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=771;
RA Zhuyoloup A.N.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74448; CAA52459.1; -.
CC PIR; S35970; S35970.
CC HAMAP; MF_00362; -; 1.
CC InterPro; IPR002363; Ribosomal_L10eub.
CC PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein.
FT INIT MET 0 0 BY SIMILARITY.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2026 MW; E20A02CD7851DB16 CRC64;

Query Match 27.1%; Score 26; DB 1; Length 19;
Best Local Similarity 31.2%; Pred. No. 4.9e+02;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 INLADRLGLSGVQEI 17
Db 2 LNLQDKQAIKVAEVS 17

RESULT 2
RL10_PROVU STANDARD; PRT; 20 AA.
AC P51411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10 (Fragment).
GN RPLJ.

```
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhyvoloup A.N.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74446; CAA52457.1; -.
CC PIR; S35978; S35978.
CC HAMAP; MF_00362; -.
CC InterPro; IPR002363; Ribosomal_L10eub.
CC PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
CC Ribosomal protein.
CC INIT MET 0 BY SIMILARITY.
CC NON TER 20 20
CC SEQUENCE 20 AA; 2083 MW; 76520A02CD7851DB CRC64;
CC
CC Query Match 27.1%; Score 26; DB 1; Length 20;
CC Best Local Similarity 31.2%; Pred. No. 5.2e+02;
CC Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 2 INLADRMGLSGVQEI 17
CC :||| : : : :
CC 2 LNLQDKQAIVAEVSEV 17
CC
CC Db
CC
CC RESULT 3
CC RL10_KLEPN STANDARD; PRT; 22 AA.
CC AC P41190;
CC DT 01-FEB-1995 (Rel. 31, Created)
CC DT 01-FEB-1995 (Rel. 31, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE 50S ribosomal protein L10 (Fragment).
CC GN RPLJ.
CC OS Klebsiella pneumoniae.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Klebsiella.
CC NCBI_TaxID=573;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=NCITC 5054;
CC Zhyvoloup A.N.;
CC Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74445; CAA52456.1; -.
CC PIR; S35975; S35975.
CC PIR; S35976; S35976.
CC HAMAP; MF_00362; -.
CC InterPro; IPR002363; Ribosomal_L10eub.
CC PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
CC Ribosomal protein.
CC INIT MET 0 BY SIMILARITY.
CC NON TER 0
```

```
FT NON TER 22 22
SQ SEQUENCE 22 AA; 2268 MW; 113D96520A02CD78 CRC64;
Query Match 27.1%; Score 26; DB 1; Length 22;
Best Local Similarity 31.2%; Pred. No. 5.7e+02;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 2 INLADRMGLSGVQEI 17
: ||| : : : :
2 LNLQDKQAIVAEVSEV 17
Db
RESULT 4
RL10_ENTCL STANDARD; PRT; 23 AA.
AC Q47608;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10 (Fragment).
GN RPLJ.
OS Enterobacter cloacae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
OC NCBI_TaxID=550;
CC [1]
CC SEQUENCE FROM N.A.
CC Zhyvoloup A.;
CC Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74444; CAA52455.1; -.
CC HAMAP; MF_00362; -.
CC InterPro; IPR002363; Ribosomal_L10eub.
CC PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
CC Ribosomal protein.
CC INIT MET 0 BY SIMILARITY.
CC NON TER 23 23
CC SEQUENCE 23 AA; 2355 MW; 39413D96520A02CD CRC64;
CC
CC Query Match 27.1%; Score 26; DB 1; Length 23;
CC Best Local Similarity 31.2%; Pred. No. 6e+02;
CC Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 2 INLADRMGLSGVQEI 17
CC : ||| : : : :
CC 2 LNLQDKQAIVAEVSEV 17
CC
CC Db
CC
CC RESULT 5
CC PSER_WHEAT STANDARD; PRT; 24 AA.
CC AC P12358;
CC DT 01-OCT-1989 (Rel. 12, Created)
CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
CC DE Photosystem II 10 kDa polypeptide (Fragment).
CC GN PSBR.
CC Triticum aestivum (Wheat).
CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
CC Triticeae; Triticum.
CC NCBI_TaxID=4565;
CC [1]
CC RN [1]
CC INIT MET 0 BY SIMILARITY.
CC NON TER 0
```

```

Qy      3 NIADMLGLGSVQ 15
      :|||:||||:
Db      1 DLAGEHGLSNVK 13

RESULT 7
YPRB SERMA
ID YPRB SERMA STANDARD; PRT; 20 AA.
AC P22581;
CT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in proB 5' region (Fragment).
OS Serratia marcescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=Str41;
RX MEDLINE=91237315; PubMed=1851803;
RA Omori K., Suzuki S., Imai Y., Komatsubara S.;
RT "Analysis of the Serratia marcescens proBA operon and feedback
RT control of proline biosynthesis.";
RL J. Gen. Microbiol. 137:509-517(1991).
CC
CC This SWISS-PROT entry is copyright. It is produced through a coll-
CC between the Swiss Institute of Bioinformatics and the EMBL out-
CC the European Bioinformatics Institute. There are no restriction
CC use by non-profit institutions as long as its content is in-
CC modified and this statement is not removed. Usage by and for c-
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90351; BAA14363.1; -
DR EMBL; X53086; CAA37253.1; -
DR PIR; C49753; C49753.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 20 AA; 2248 MW; 4DD777735276674 CRC64;

Query Match 26.0%; Score 25; DB 1; Length 20;
Best Local Similarity 41.7%; Pred. No. 7.6e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0;

Qy      9 IGLSGVQEIKEQ 20
      ||||:|
Db      3 IGLPEADDFKEK 14

RESULT 8
ID FLB3 TREHY
FLB3 TREHY STANDARD; PRT; 25 AA.
AC P80161;
CT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar filament core protein flab3 (32 kDa core protein)
DE (Fragment).
DE DE
GN FLAB3.
OS Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC Bacteria; Spirochaetes; Spirochaetales; Brachyspiraceae; Brachyspi-
OX NCBI_TaxID=159;
RN [1]
RP SEQUENCE.
RC STRAIN=C5;
RX MEDLINE=931139764; PubMed=1487733;
RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
RA van der Zeijst B.A.M., Kusters J.G.;
RT "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae
RT composed of two sheath proteins and three core proteins.";
RL J. Gen. Microbiol. 138:2697-2706(1992).

```

```

CC -!- FUNCTION: Component of the core of the flagella.
CC -!- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF TWO
CC SHEATH PROTEINS, FLAA1 (44 kDa) AND FLAA2 (35 kDa) AROUND A CORE
CC THAT CONTAINS THREE PROTEINS FLAB1 (37 kDa), FLAB2 (34 kDa) AND
CC FLAB3 (32 kDa).
CC -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
CC -!- SIMILARITY: Belongs to the bacterial flagellin family.
DR PIR: D47689; D47689.
KW Flagellum; Periplasmic.
FT NON TR 25 25
SQ SEQUENCE 25 AA; 2746 MW; 1AA0ACAV722EE3D0A CRC64;

Query Match 26.0%; Score 25; DB 1; Length 25;
Best Local Similarity 55.6%; Pred. No. 9.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 ADRMLGLSG 13
DB 12 ANRQLNLTG 20

RESULT 9
Y523 BORBU STANDARD; PRT; 30 AA.
AC O51473;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BB0523.
GN BB0523.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
SEQUENCE FROM N.A.
STRAIN=ATCC 35210 / B31;
MEDLINE=98065943; PubMed=9403685;
Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lachigra R., White O., Ketchum K.A., Dodson R., Hickson E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Hanson M.,
Peterson J., Kervilave A.R., Quackenbush J., Salzberg S., Weidman J.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
Utterback T., Wattley L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;
"Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi.";
Nature 390:580-586(1997).

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AE001154; AAC66894.1; --
PIR: B70165; B70165.
TIGR: BB0523; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 30 AA; 3431 MW; DE437B15D2A48AA8 CRC64;

Query Match 26.0%; Score 25; DB 1; Length 30;
Best Local Similarity 25.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 9 LGLSGVQBIKEQ 20
DB 1 MGFNGILAKQR 12

RESULT 10

```

```

GALA ONCMY STANDARD; PRT; 29 AA.
ID P47213; 1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Galanin.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
SEQUENCE.
TISSUE=Stomach; PubMed=7532194;
RX MEDLINE=95164756;
RA Anglade I., Wang Y., Jensen J., Tramu G., Kah O., Conlon J.M.;
RT "Characterization of trout galanin and its distribution in trout
brain and pituitary.";
RL J. Comp. Neurol. 350:63-74(1994).
CC -!- FUNCTION: Contracts smooth muscle of the gastrointestinal and
CC genitourinary tract, regulates growth hormone release, modulates
CC insulin release, and may be involved in the control of adrenal
CC secretion.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the galanin family.
CC InterPro: IPR008174; Galanin.
DR Pfam: PF01296; Galanin; 1.
DR PRINTS: PR00273; GALANIN.
DR PRODOM: PD005962; Galanin; 1.
DR PROSITE: PS00861; GALANIN; 1.
KW Hormone; Neuropeptide; Amidation.
FT MOD RES 29 29
SQ SEQUENCE 29 AA; 3044 MW; 73C37190403FA349 CRC64;

Query Match 25.0%; Score 24; DB 1; Length 29;
Best Local Similarity 38.5%; Pred. No. 1.6e+03;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 INLADRMGLSGV 14
DB 4 LNSAGYLGLPHGI 16

RESULT 11
CD11_LITXA STANDARD; PRT; 12 AA.
AC P56245; P81253;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Caeridin 1.1/1.2/1.3.
OS Litoria xanthomera (Orange-thighed frog),
OS Litoria splendida (Magnificent tree frog),
OS Litoria chloris (Centralian tree frog), and
OS Litoria chloris (Blue-thighed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaidea;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=79697; 30345, 39405, 86064;
RN [1]
SEQUENCE, AND MASS SPECTROMETRY.
RP SPECIES=L.xanthomera;
RX MEDLINE=97374000; PubMed=9230483;
RA Steinbörner S.T., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.,
RA Ramsay S.L.;
RT "New caerin antibacterial peptides from the skin glands of the
RL Australian tree frog Litoria xanthomera.";
RL J. Pept. Sci. 3:181-185(1997).
RN [2]
SEQUENCE, AND MASS SPECTROMETRY.
RP SPECIES=L.splendida; TISSUE=Parotoid gland;
RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;

```


RT "Peptides from Australian frogs. Structures of the caerins and
 RT Caeridin 1 from Litoria splendida.";
 RL J. Chem. Soc. Perkin Trans. 1:3173-3178(1992).
 RN [3]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC SPECIES=L.gilleni; TISSUE=Parotoid gland;
 RA Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;
 RT "Peptides from Australian frogs. The structures of the caerins and
 RT caeridins from Litoria gilleni.";
 RL J. Chem. Res. 139:937-961(1993).
 RN [4]
 RP SEQUENCE, AND MASS SPECTROMETRY (CAERIDINS 1.1; 1.2 AND 1.3).
 RC SPECIES=L.gilleni;
 RA Waugh R.J., Steinborner S.T., Bowie J.H., Wallace J.C., Tyler M.J.,
 RA Hu P., Gross M.L.;
 RT "Two isomeric alpha and beta aspartyl dodecapeptides and their
 RT cyclic amino succinyl analogue from the Australian tree frog
 RT Litoria gilleni.";
 RL Aust. J. Chem. 48:1981-1987(1995).
 RN [5]
 RP SEQUENCE.
 RC SPECIES=L.chloris; TISSUE=Skin secretion;
 RX MEDLINE=98175802; PubMed=9516047;
 RA Steinborner S.T., Currie G.J., Bowie J.H., Wallace J.C., Tyler M.J.;
 RT "New antibiotic caerin 1 peptides from the skin secretion of the
 RT Australian tree frog Litoria chloris. Comparison of the activities of
 RT the caerin 1 peptides from the genus Litoria.";
 RL J. Pept. Res. 51:121-126(1998).
 CC -!- FUNCTION: Caeridins show neither neuropeptide activity nor
 CC antibiotic activity.
 CC -!- TISSUE SPECIFICITY: Secreted by the skin dorsal glands in
 CC L.xanthomera or specifically by the skin parotoid and/or rostral
 CC glands in L.splendida and L.gilleni.
 CC -!- PTM: ISOMERIZATION ALPHA-BETA OF THE ASP-4 RESIDUE IN CAERIDIN
 CC 1.2; A CYCLIC SUCCINIMIDE MAY BE FORMED BETWEEN ASP-4 AND GLY-5
 CC RESIDUES IN CAERIDIN 1.3.
 CC -!- MASS SPECTROMETRY: MW=1140; METHOD=FAB.
 CC -!- MISCELLANEOUS: The results for mass spectrometry of caeridin 1.1
 CC is 1140 in Ref.1, Ref.3 and Ref.4, while in Ref.2 it is 1139.
 KW Amphibian defense peptide; Amidation.
 FT MOD_RES 12 12 AMIDATION.
 SQ SEQUENCE 12 AA; 1141 MW; 2822551A33772728 CRC64;
 Query Match 24.0%; Score 23; DB 1; Length 12;
 Best Local Similarity 45.5%; Pred. No. 9.5e+02;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 4 LADRLMLGLSGV 14
 DB 2 LLDGLGLTGL 12
 RESULT 12
 FMO1 CAVPO
 ID FMO1 CAVPO STANDARD; PRT; 19 AA.
 AC P49328;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Dimethylaniline monooxygenase [N-oxide forming] 1 (EC 1.14.13.8)
 DE (Hepatic flavin-containing monooxygenase 1) (FMO 1) (Dimethylaniline
 DE oxidase 1) (Fragment).
 GN FMO1 OR FMO-1.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=90314396; PubMed=2369122;
 RA Yamada H., Yuno K., Oguri K., Yoshimura H.;
 RT "Multiplicity of liver microsomal flavin-containing monooxygenase in

RT the guinea pig: its purification and characterization.";
 RL Arch. Biochem. Biophys. 280:305-312(1990).
 CC -!- FUNCTION: This protein is involved in the oxidative metabolism of
 CC a variety of xenobiotics such as drugs and pesticides.
 CC -!- CATALYTIC ACTIVITY: N,N-dimethylaniline + NADPH + O(2) = N,N-
 CC dimethylaniline N-oxide + NADP(+) + H(2)O.
 CC -!- COFACTOR: FAD.
 CC -!- SUBCELLULAR LOCATION: Microsomal.
 CC -!- TISSUE SPECIFICITY: Liver.
 CC -!- SIMILARITY: Belongs to the FMO family.
 DR PIR; S10747; S10747.
 KW Oxidoreductase; Monooxygenase; NADP; Flavoprotein; FAD; Microsome;
 KW Transmembrane; Multigene family.
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 1826 MW; 324DDEB75CA05522 CRC64;
 Query Match 24.0%; Score 23; DB 1; Length 19;
 Best Local Similarity 44.4%; Pred. No. 1.5e+03;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 10 GLSGVQEI 18
 DB 10 GVSGLASIR 18
 RESULT 13
 UP27 UPEMJ
 ID UP27 UPEMJ STANDARD; PRT; 19 AA.
 AC P82039;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Uperin 2.7.
 OS Uperoleia mjobergii (Australian toadlet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylaidea; Myobatrachidae;
 OC Myobatrachinae; Uperoleia.
 OX NCBI_TaxID=104954;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Skin secretion;
 RA Bradford A.M., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "New antibiotic uperin peptides from the dorsal glands of the
 RT Australian toadlet Uperoleia mjobergii.";
 RL Aust. J. Chem. 49:1325-1331(1996).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
 CC -!- MASS SPECTROMETRY: MW=1948; METHOD=FAB.
 KW Amphibian defense peptide.
 SQ SEQUENCE 19 AA; 1949 MW; 24E4F83A6BA35F21 CRC64;
 Query Match 24.0%; Score 23; DB 1; Length 19;
 Best Local Similarity 25.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 8; Mismatches 2; Indels 2; Gaps 1;
 QY 2 INLADRLMLGLSGVQEI 17
 DB 3 IDIAKKLVG--GIRNV 16
 RESULT 14
 NDP2 ARTSK
 ID NDP2 ARTSK STANDARD; PRT; 23 AA.
 AC P83576;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE NADP phosphatase II (EC 3.1.3.-) (Fragment).
 OS Arthrobacter sp. (strain KM).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcineae; Micrococcaceae; Arthrobacter.
 OX NCBI_TaxID=184230;
 RN [1]

Job time : 5.33333 secs

RP SEQUENCE, SUBUNIT, AND SUBCELLULAR LOCATION.
RA Kawai S., Mori S., Mukai T., Murata K.;
RT "Properties of NADP phosphatase I and II of *Arthrobacter* sp. strain
KM.";
RL Bur. J. Biochem. 0:0-0(2003).
CC -|- SUBUNIT: Homodimer.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- MISCELLANEOUS: Maximal activity at pH 7-8.
DR InterPro; IPR000437; Prok_lipoprot_S.
KW Hydrolase.
FT NON TER 23 23
SQ SEQUENCE 23 AA; 2453 MW; 19547F56E0655886 CRC64;

Query Match 24.0%; Score 23; DB 1; Length 23;
Best Local Similarity 44.4%; Pred. No. 1.8e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KINLADRML 9
DB 7 XANLGNQMI 15
||| :||
||| :||

RESULT 15
CR22 LITGI STANDARD; PRT; 25 AA.
AC P56234;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Caerin 2.2 [Contains: Caerin 2.2.1].
OS Litoria gilleni (Centralian tree frog), and
OS Litoria caerulea (Green tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=39405, 30344;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY OF CAERIN 2.2.
RC SPECIES=L.caerulea; TISSUE=Parotoid gland;
RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins from
Litoria caerulea.";
RL J. Chem. Res. 138:910-936(1993).
RN [2]
RP SEQUENCE, AND MASS SPECTROMETRY OF CAERIN 2.2.1.
RC SPECIES=L.gilleni; TISSUE=Parotoid gland;
RA Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins and
caeridins from Litoria gilleni.";
RL J. Chem. Res. 139:937-961(1993).
CC -|- FUNCTION: ANTIMICROBIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL
CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Secreted by the skin parotoid and/or rostral
glands.
CC -|- MASS SPECTROMETRY: MW=2464; METHOD=FAB; RANGE=1-25.
CC -|- MASS SPECTROMETRY: MW=1695; METHOD=FAB; RANGE=9-25.
KW Amphibian defense peptide; Antibiotic.
FT PEPTIDE 1 25 CAERIN 2.2.
FT PEPTIDE 9 25 CAERIN 2.2.1.
SQ SEQUENCE 25 AA; 2466 MW; DDCA9BC5D49186B8 CRC64;

Query Match 24.0%; Score 23; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 8; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 7 RMLG--LSGVQRIKEQ 20
DB 8 RALGGILLADVKSKEQ 23
||| :|||
||| :|||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:31 ; Search time 25.7333 Seconds
(without alignments)
245.221 Million cell updates/sec

Title: US-09-171-432A-46
Perfect score: 96
Sequence: 1 KINLADRLMLGLSGVQEIKEQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 18020

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	31	32.3	30	8	Q9T2P9	Q9c2p9 narcissus p
2	30	31.2	17	2	Q9X3I3	Q9x3i3 prochloroco
3	29	30.2	26	4	Q9BRY8	Q9bry8 homo sapien
4	27	28.1	19	2	Q7WV3	Q7wv3 streptococc
5	27	28.1	22	10	Q9SAU4	Q9sau4 arabidopsis
6	27	28.1	23	12	Q68983	Q68983 suid herpes
7	27	28.1	25	5	Q25183	Q25183 hydra atten
8	27	28.1	26	2	Q93HQ7	Q93hq7 streptococc
9	27	28.1	29	4	Q9UM88	Q9um88 homo sapien
10	26.5	27.6	27	5	Q61676	Q61676 drosophila
11	26	27.1	14	11	Q51864	Q51864 mus musculus
12	26	27.1	16	10	Q9AUB0	Q9aub0 olimarabido
13	26	27.1	16	10	Q9AUB1	Q9aub1 arabidopsis
14	26	27.1	24	7	O78181	O78181 homo sapien
15	26	27.1	24	12	Q9PXB7	Q9pxb7 classical s
16	26	27.1	25	11	Q63990	Q63990 rattus norv

17	26	27.1	27	5	P91901	P91901 beroe ovata
18	26	27.1	27	5	Q9BM49	Q9bm49 moniliformi
19	26	27.1	27	13	Q90297	Q90297 carassius a
20	26	27.1	28	2	Q47368	Q47368 escherichia
21	26	27.1	28	2	Q47360	Q47360 escherichia
22	26	27.1	28	2	Q47366	Q47366 escherichia
23	26	27.1	29	2	Q47389	Q47389 escherichia
24	25.5	26.6	26	5	Q17855	Q17855 caenorhabdi
25	25	26.0	14	13	O73591	O73591 gallus gall
26	25	26.0	19	2	O8GL20	O8gl20 borrelia bu
27	25	26.0	20	3	P82288	P82288 acrononium
28	25	26.0	21	2	Q9X3I1	Q9x3i1 prochloroco
29	25	26.0	21	4	Q9UCC3	Q9ucc3 homo sapien
30	25	26.0	24	5	Q9U9B7	Q9u9b7 ceratitis c
31	25	26.0	24	13	Q7T251	Q7t251 brachydanio
32	25	26.0	25	5	Q94681	Q94681 polyandroca
33	25	26.0	26	17	O8ZV3	O8z2v3 pyrobaculum
34	25	26.0	27	5	Q25482	Q25482 metridium s
35	25	26.0	27	5	O02602	O02602 beroe ovata
36	25	26.0	27	12	Q9QI46	Q9qi46 hepatitis c
37	25	26.0	27	12	Q9QI39	Q9qi39 hepatitis c
38	25	26.0	28	2	Q47354	Q47354 escherichia
39	25	26.0	28	2	Q47359	Q47359 escherichia
40	25	26.0	28	2	Q47358	Q47358 escherichia
41	25	26.0	28	2	Q47370	Q47370 escherichia
42	25	26.0	28	2	Q47352	Q47352 escherichia
43	25	26.0	28	2	Q47369	Q47369 escherichia
44	25	26.0	28	2	Q47356	Q47356 escherichia
45	25	26.0	28	2	Q47363	Q47363 escherichia

ALIGNMENTS

RESULT 1

Q9T2P9 PRELIMINARY; PRT; 30 AA.
ID Q9T2P9
AC Q9T2P9
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Heat shock protein 60 (Fragment)
OS Narcissus pseudonarcissus (Daffodil)
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Narcissus.
OX NCBI_TaxID=39639;
RN [1]_TaxID=39639;
RP SEQUENCE.
RX MEDLINE=96291727; PubMed=8754688;
RA Bonk M., Tadros M., Vandekerckhove J., Al-Babili S., Beyer P.;
RT "Purification and characterization of chaperonin 60 and heat-shock
RT protein 70 from chromoplasts of Narcissus pseudonarcissus.";
RL Plant Physiol. 111:931-939(1996).
DR HSP; P61399; IAOB.
DR InterPro; IPR008950; GroEL-ATPase.
SQ SEQUENCE 30 AA; 3233 MW; AF5AF69899CE2851 CRC64;

Query Match 32.3%; Score 31; DB 8; Length 30;
Best Local Similarity 36.8%; Pred. No. 8.2e+02;
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 1 KINLADRLMLGLSGVQEIKE 19
Db 6 KFGVEARALMLRGVEELAD 24

RESULT 2

Q9X3I3 PRELIMINARY; PRT; 17 AA.
ID Q9X3I3
AC Q9X3I3
DT 01-NOV-1999 (TREMBLrel. 12, Created)

```

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Cytochrome b (Fragment).
GN PETB.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630 (1998).
DR EMBL; AF070176; AAD20791.1; -.
FT NON TER 1
SQ SEQUENCE 17 AA; 1953 MW; 630843039ADD51B4 CRC64;

Query Match 31.2%; Score 30; DB 2; Length 17;
Best Local Similarity 60.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 LADRMGLSG 13
DB 6 LXDRKQGISG 15

RESULT 3
Q98Y8 PRELIMINARY; PRT; 26 AA.
AC Q98Y8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Straussberg R.;
RC TISSUE=Colon;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005853; AAH05893.1; -.
KW Hypothetical protein.
SQ SEQUENCE 26 AA; 2723 MW; E5CC438009845B82 CRC64;

Query Match 30.2%; Score 29; DB 4; Length 26;
Best Local Similarity 38.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 INLADRMGLSGV 14
DB 3 LSLASQSAGITGV 15

RESULT 4
Q7WV3 PRELIMINARY; PRT; 19 AA.
AC Q7WV3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ALIA (Fragment).
GN ALIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type 5;

```

```

RA Xie X., Fletcher L., Farley J., Russell D., Adegbola R., Murphy E.,
RA Zagursky R.J.;
RT "A new cap5 operon in Streptococcus pneumoniae located between dexB
RT and aliA.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY336008; AAP94622.1; -.
FT NON TER 19
SQ SEQUENCE 19 AA; 1947 MW; E6722E0D350A6E51 CRC64;

Query Match 28.1%; Score 27; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 2.3e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 ADMRLGLSGV 14
DB 4 SSRLLALAGV 13

RESULT 5
Q9SAU4 PRELIMINARY; PRT; 22 AA.
AC Q9SAU4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Chalcone synthase (Fragment).
GN CHS.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Trezzini G.F., Horrichs A., Somssich I.E.;
RC MEDLINE=93144713; PubMed=8425063;
RX "Isolation of putative defense-related genes from Arabidopsis thaliana
RX and expression in fungal elicitor-treated cells.";
RL Plant Mol. Biol. 21:385-389 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Hartmann U., Valentine W.J., Christie J.M., Hays J., Jenkins G.I.,
RA Weisstaar B.;
RT "Identification of UV/blue light-response elements in the Arabidopsis
RT thaliana chalcone synthase promoter using a homologous protoplast
RT transient expression system.";
RL Plant Mol. Biol. 36:741-754 (1998).
DR EMBL; AF012810; AAC17466.2; -.
FT NON TER 22
SQ SEQUENCE 22 AA; 2275 MW; 0D31732B4630BD74 CRC64;

Query Match 28.1%; Score 27; DB 10; Length 22;
Best Local Similarity 41.7%; Pred. No. 2.7e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 MGLSGVQEIKE 19
DB 3 MAGASSLDEIRQ 14

RESULT 6
Q68983 PRELIMINARY; PRT; 23 AA.
ID Q68983
AC Q68983;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycoprotein C precursor (Fragment).
GN GC.
OS Suid herpesvirus 1.

```

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10345;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Becker;
RX MEDLINE=95081163; PubMed=7989378;
RA Tomillo M., Wilkinson K.S., Ryan P.;
RT "Can a signal sequence become too hydrophobic?";
RL J. Biol. Chem. 269:32016-32021(1994).
DR EMBL; L36970; AAA79967.1; -.
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2248 MW; B708473E8146D98F CRC64;

Query Match 28.1%; Score 27; DB 12; Length 23;
Best Local Similarity 66.7%; Pred. No. 2.9e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NLADRMIGL 11
Db :||| |||
3 SLADAMLAL 11

RESULT 7
Q25183
ID Q25183 PRELIMINARY; PRT; 25 AA.
AC Q25183;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE CnHv2 protein (Fragment).
GN CnHv2
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Chnidaria; Hydrozoa; Hydroids; Anthomedusae;
OC Hydrozoa; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93321609; PubMed=8101168;
RA Shenk M.A., Bode H.R., Steele R.E.;
RT "Expression of Cnox-2, a HOM/HOX homeobox gene in hydra, is correlated with axial pattern formation.";
RL Development 117:657-667(1993).
DR EMBL; M62871; AAX29208.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3144 MW; 90936F858F24D93C CRC64;

Query Match 28.1%; Score 27; DB 5; Length 25;
Best Local Similarity 44.4%; Pred. No. 3.1e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KINLADRM 9
Db :||| :|||
16 KINLSERQI 24

RESULT 8
Q93HQ7
ID Q93HQ7 PRELIMINARY; PRT; 26 AA.
AC Q93HQ7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Homologue of yloU of B. subtilis (Fragment).
GN HLO.

OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BL(T);
RA Karaya K., Taketo A.;
RT "Streptococcal DNA regions related to streptolysin S (SLS) formation";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB069726; BAB64350.1; -.
FT NON_TER 26 26
SQ SEQUENCE 26 AA; 2676 MW; 9CB5232B399DE056 CRC64;

Query Match 28.1%; Score 27; DB 2; Length 26;
Best Local Similarity 50.0%; Pred. No. 3.3e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KINLADRMIGLS 12
Db :||| :|||
4 KINTKDGIELS 15

RESULT 9
Q9UM88
ID Q9UM88 PRELIMINARY; PRT; 29 AA.
AC Q9UM88;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Beta 2-microglobulin protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92145672; PubMed=1737380;
RA Gattoni-Celli S., Kirsch K., Timpone R., Isselbacher K.J.;
RT "Beta 2-microglobulin gene is mutated in a human colon cancer cell line (HCT) deficient in the expression of HLA class I antigens on the cell surface.";
RL Cancer Res. 52:1201-1204(1992).
DR EMBL; S82300; AAD14388.1; -.
SQ SEQUENCE 29 AA; 3066 MW; 3F4C8934779CF041 CRC64;

Query Match 28.1%; Score 27; DB 4; Length 29;
Best Local Similarity 45.5%; Pred. No. 3.7e+03;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 8 MLGLSGVQEI 18
Db :||| :|||
12 LLSLSGLEAIQ 22

RESULT 10
Q61676
ID Q61676 PRELIMINARY; PRT; 27 AA.
AC Q61676;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE HOX11-D125 protein (D125) (Fragment).
GN HOX11-D125 OR D125.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281593; PubMed=8099440;

Dear T.N., Sanchez-Garcia I., Rabbitts T.H.;
RT "The HOX11 gene encodes a DNA-binding nuclear transcription factor
belonging to a distinct family of homeobox genes."
RL Proc. Natl. Acad. Sci. U.S.A. 90:4431-4435(1993).
DR EMBL: L08618; AAA28612.1; --
DR FlyBase: FBgn0010394; Hox11-D125.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR ProDom: PD000010; Homeobox; 1.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 2937 MW; CABBC2F876F2B3A3 CRC64;
Query Match 27.6%; Score 26.5; DB 5; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.1e+03;
Matches 9; Conservative 0; Mismatches 4; Indels 5; Gaps 1;
QY 5 ADR-----MLGLSGVQEI 17
Db 9 ADRDEIAAASLGLSNAQVI 26
RESULT 11
Q61864 PRELIMINARY; PRT; 14 AA.
AC Q61864;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integral membrane protein (Fragment).
GN J-ALPHA TA80.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=92165347; PubMed=1371499;
RA Nakajima P.B., Di Vincenzo J.P., Jameson S.C., Gascoigne N.R.J.;
RT "Chromosome 14 in B10.A (18R) mice is recombinant and includes Tcra-V
alleles."
RL Immunogenetics 35:190-198(1992).
DR EMBL: M55635; AAA39704.1; --
DR PIR: I68824; I68824.
FT NON_TER 1
SQ SEQUENCE 14 AA; 1492 MW; CE1CFD0FAC21A379 CRC64;
Query Match 27.1%; Score 26; DB 11; Length 14;
Best Local Similarity 46.2%; Pred. No. 2.5e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 5 ADRMLGLSGVQEI 17
Db 2 ADRLTFFGKGTOOI 14
RESULT 12
Q9AUB0 PRELIMINARY; PRT; 16 AA.
AC Q9AUB0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Chalcone synthase (Fragment).
OS Olinarabidopsis pumila (Arabidopsis pumila).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Olinarabidopsis.
OX NCBI_TaxID=74718;

SEQUENCE FROM N.A.
RP Koch M., Kroymann J., Haubold B., Weisshaar B., Mitchell-Olds T.;
RA "Phylogenetic analysis of promoter sequences from cruciferous
RT plants."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF248989; AAK31924.1; --
DR NON_TER 16
SQ SEQUENCE 16 AA; 1707 MW; 7D745C76CB016176 CRC64;
Query Match 27.1%; Score 26; DB 10; Length 16;
Best Local Similarity 41.7%; Pred. No. 2.8e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 8 MLGLSGVQEIKE 19
Db 3 MAGASSLDEIRK 14
RESULT 13
Q9AUB1 PRELIMINARY; PRT; 16 AA.
AC Q9AUB1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Chalcone synthase (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Koch M., Kroymann J., Haubold B., Weisshaar B., Mitchell-Olds T.;
RT "Phylogenetic analysis of promoter sequences from cruciferous
RT plants."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF248988; AAK31923.1; --
DR NON_TER 16
SQ SEQUENCE 16 AA; 1707 MW; 7D745C76CB016176 CRC64;
Query Match 27.1%; Score 26; DB 10; Length 16;
Best Local Similarity 41.7%; Pred. No. 2.8e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 8 MLGLSGVQEIKE 19
Db 3 MAGASSLDEIRK 14
RESULT 14
Q78181 PRELIMINARY; PRT; 24 AA.
AC Q78181;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Minor histocompatibility antigen HA-1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilke M., Pool J., Den Haan J., Goulmy E.;
RT "Genomic identification of the minor histocompatibility antigen HA-1
RT locus by allele-specific PCR."
RL Tissue Antigens 52:0-0(1998).
DR EMBL: AF092537; AAC64140.1; --
DR NON_TER 1
FT NON_TER 24
SQ SEQUENCE 24 AA; 2713 MW; 39ADF9FFB89F1280 CRC64;

Query Match 27.1%; Score 26; DB 7; Length 24;
 Best Local Similarity 37.5%; Pred. No. 4.4e+03;
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 LADRLMLGLSGVQEIKE 19
 ||| |::||
 Db 4 LADVAFAEGLKLE 19

RESULT 15
 Q9FAB7
 ID Q9PXE7 PRELIMINARY; PRT; 24 AA.
 AC Q9PXE7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Glycoprotein E0 (Fragment).
 OS Classical swine fever virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Pestivirus.
 OX NCBI_TaxID=11096;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93267778; PubMed=8388499;
 RA Rumenapf T., Unger G., Strauss J.H., Thiel H.J.;
 RT "Processing of the envelope glycoproteins of pestiviruses."
 RL J. Virol. 67:3288-3294(1993).
 SQ SEQUENCE 24 AA; 2752 MW; A48D322C89550658 CRC64;

Query Match 27.1%; Score 26; DB 12; Length 24;
 Best Local Similarity 46.2%; Pred. No. 4.4e+03;
 Matches 6; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 3 NLADRLMLGLSGVQ 15
 ||| |::||
 Db 7 NLSDN--GTNGIQ 17

Search completed: July 15, 2004, 14:51:03
 Job time : 27.7333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:47:02 ; Search time 12.4 Seconds
(without alignments)
83.268 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRMGLSGVQEIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 204689

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	35.4	28	1	US-07-977-630-45
2	31	32.3	25	1	US-08-468-674B-15
3	31	32.3	25	1	US-08-780-571-15
4	31	32.3	25	4	US-09-324-217B-17
5	30	31.2	28	3	US-09-295-186-3
6	30	31.2	30	3	US-09-043-632-79
7	29	30.2	22	3	US-08-940-095-173
8	29	30.2	22	3	US-08-940-093-173
9	29	30.2	22	3	US-08-940-096-173
10	29	30.2	22	3	US-09-465-719-173
11	29	30.2	22	4	US-09-453-605-173
12	29	30.2	22	4	US-09-453-838-173
13	29	30.2	22	4	US-08-940-136-173
14	29	30.2	22	4	US-09-453-841-173
15	29	30.2	22	4	US-09-453-833-173
16	29	30.2	22	4	US-09-453-826-173
17	28	29.2	11	3	US-09-215-966-20
18	28	29.2	19	3	US-08-779-764A-41
19	28	29.2	19	4	US-09-563-456-41
20	28	29.2	23	1	US-08-485-588-10
21	28	29.2	23	1	US-08-484-565-10
22	28	29.2	23	2	US-08-480-751-10
23	28	29.2	23	2	US-08-943-986-10
24	28	29.2	23	3	US-08-353-784-10
25	28	29.2	23	3	US-08-484-719B-10
26	28	29.2	23	4	US-08-484-159-10
27	28	29.2	23	4	US-09-687-477-19

28 28 29.2 23 4 US-09-687-476-19 Sequence 19, Appl
29 28 29.2 23 4 US-09-687-372-19 Sequence 19, Appl
30 28 29.2 23 4 US-09-975-553-19 Sequence 19, Appl
31 28 29.2 23 4 US-10-270-795-19 Sequence 19, Appl
32 28 29.2 23 4 US-10-270-876-19 Sequence 19, Appl
33 27 28.1 15 1 US-07-985-691-14 Sequence 14, Appl
34 27 28.1 21 1 US-08-290-448A-40 Sequence 40, Appl
35 27 28.1 21 1 US-08-290-448A-51 Sequence 51, Appl
36 27 28.1 21 1 US-08-290-448A-40 Sequence 40, Appl
37 27 28.1 21 1 US-08-290-448A-51 Sequence 51, Appl
38 27 28.1 21 1 US-08-175-069A-40 Sequence 40, Appl
39 27 28.1 21 1 US-08-175-069A-51 Sequence 51, Appl
40 27 28.1 21 4 US-08-461-939B-40 Sequence 40, Appl
41 27 28.1 21 4 US-08-461-939B-51 Sequence 51, Appl
42 27 28.1 21 4 US-08-464-000-40 Sequence 40, Appl
43 27 28.1 21 4 US-08-464-000-51 Sequence 51, Appl
44 27 28.1 27 3 US-08-360-107A-87 Sequence 87, Appl
45 26.5 27.6 15 6 5223254-3 Patent No. 5223254

ALIGNMENTS

RESULT 1

US-07-977-630-45

; Sequence 45, Application US/07977630

; Patent No. 5583038

; GENERAL INFORMATION:

; APPLICANT: Stover, Charles K.

; TITLE OF INVENTION: BACTERIAL EXPRESSION VECTORS CONTAINING

; NUMBER OF INVENTION: DNA ENCODING SECRETION SIGNALS OF LIPOPROTEINS

; NUMBER OF SEQUENCES: 84

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,

; ADDRESSER: Stewart & Olstein

; STREET: 6 Becker Farm Road

; CITY: Roseland

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette

; COMPUTER: IBM

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/977,630

; FILING DATE: No. 5583038ember 17, 1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Herron, Charles J.

; REGISTRATION NUMBER: 28,019

; REFERENCE/DOCKET NUMBER: 469201-174

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 45:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 28 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-07-977-630-45

Query Match 35.4%; Score 34; DB 1; Length 28;

Best Local Similarity 46.2%; Pred. No. 19;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 LADRMGLSGVQEQ 16

Db 15 LVERMISVSGVAD 27

```
RESULT 2
US-08-468-674B-15
; Sequence 15, Application US/08468674B
; Patent No. 5639642
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas B
; APPLICANT: Vad, Knud
; TITLE OF INVENTION: SYNTHETIC LEADERS PEPTIDE SEQUENCES
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 56396420 No. 5639642disk of No. 5639642th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,674B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/282,852
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4085.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-468-674B-15

Query Match 32.3%; Score 31; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 INLADRMGLSG 13
Db 13 VNLMDLAVGLPG 24

RESULT 3
US-08-780-571-15
; Sequence 15, Application US/08780571
; Patent No. 5795746
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas B
; APPLICANT: Vad, Knud
; TITLE OF INVENTION: SYNTHETIC LEADERS PEPTIDE SEQUENCES
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57957460 No. 5795746disk of No. 5795746th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780,571
; FILING DATE: 08-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/468,674
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: US 08/282,852
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4085.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-780-571-15

Query Match 32.3%; Score 31; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 INLADRMGLSG 13
Db 13 VNLMDLAVGLPG 24

RESULT 4
US-09-324-217B-17
; Sequence 17, Application US/09324217B
; Patent No. 6500645
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas Boeglum
; APPLICANT: Balschmidt, Per
; APPLICANT: Pettersson, Annette Frost
; APPLICANT: Vad, Knud
; APPLICANT: Brandt, Jakob
; APPLICANT: Havelund, Svend
; TITLE OF INVENTION: N-Terminally Extended Proteins Expressed
; TITLE OF INVENTION: In Yeast
; FILE REFERENCE: 5930.500-US
; CURRENT APPLICATION NUMBER: US/09/324,217B
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/036,758
; PRIOR FILING DATE: 1997-01-28
; PRIOR APPLICATION NUMBER: 0712/94
; PRIOR FILING DATE: 1994-06-17
; PRIOR APPLICATION NUMBER: 1449/95
; PRIOR FILING DATE: 1995-12-20
; PRIOR APPLICATION NUMBER: 1482/96
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Variation
US-09-324-217B-17

Query Match 32.3%; Score 31; DB 4; Length 25;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 INLADRMGLSG 13
Db 13 VNLMDLAVGLPG 24
```

Db 13 VNLMDLAVGLPG 24

RESULT 5

US-09-295-186-3

; Sequence 3, Application US/09295186B

; Patent No. 6127137

; GENERAL INFORMATION:

; APPLICANT: Hasida, Miyoko

; APPLICANT: Tsutsumi, No. 61271371ko

; APPLICANT: Halkier, Torben

; APPLICANT: Stringer, Mary Ann

; TITLE OF INVENTION: An Acidic Phospholipase, Production, and

; FILE OF INVENTION: Methods of Using Thereof (As Amended)

; FILE REFERENCE: 4953.204-US

; CURRENT APPLICATION NUMBER: US/09/295,186B

; CURRENT FILING DATE: 1999-04-20

; PRIOR APPLICATION NUMBER: 1215/96

; PRIOR FILING DATE: 1996-10-31

; PRIOR APPLICATION NUMBER: PCT/DK97/00490

; PRIOR FILING DATE: 1997-10-30

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 28

; TYPE: PRT

; ORGANISM: Hypozyma sp. CBS 648.91

; FEATURE:

; OTHER INFORMATION: Xaa at position 4 is any amino acid

US-09-295-186-3

Query Match 31.2%; Score 30; DB 3; Length 28;

Best Local Similarity 50.0%; Pred. No. 1e+02;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 INLDRMLGLSG 13

Db 16 LQLAEYIAGLSG 27

RESULT 6

US-09-045-632-79

; Sequence 79, Application US/09045632

; Patent No. 6001575

; GENERAL INFORMATION:

; APPLICANT: Huganir, Richard L.

; APPLICANT: Dong, Hualing

; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND

; FILE OF INVENTION: GRIP-RELATED MOLECULES

; NUMBER OF SEQUENCES: 105

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

; STREET: 130 Water Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/045,632

; FILING DATE: 19-MAR-1998

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/041,016

; FILING DATE: 19-MAR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Corleiss, Peter F.

; REGISTRATION NUMBER: 33,860

; REFERENCE/DOCKET NUMBER: 48147/1699-CIP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 79:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 30 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-045-632-79

Query Match 31.2%; Score 30; DB 3; Length 30;

Best Local Similarity 23.1%; Pred. No. 1.1e+02;

Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 INLDRMLGLSGV 14

Db 17 LQIGDRVMAINGI 29

RESULT 7

US-08-940-095-173

; Sequence 173, Application US/08940095

; Patent No. 6004925

; GENERAL INFORMATION:

; APPLICANT: Dasseux, Jean-Louis

; APPLICANT: Sekul, Renate

; APPLICANT: Buttner, Klaus

; APPLICANT: Cornut, Isabelle

; APPLICANT: Metz, Gunther

; APPLICANT: Dufourcq, Jean

; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS

; FILE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS

; NUMBER OF SEQUENCES: 258

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036-2811

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/940,095

; FILING DATE: 29-SEP-1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 009196-0004-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-493-4935

; TELEFAX: 650-493-5556

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 173:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: No. 6004925e

US-08-940-095-173

Query Match 30.2%; Score 29; DB 3; Length 22;

Best Local Similarity 66.7%; Pred. No. 1.1e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NLADRMGL 11
|||:|
Db 8 NLLDRLLDL 16

RESULT 8
US-08-940-093-173
; Sequence 173, Application US/08940093
; Patent No. 6037323
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940.093
; FILING DATE: 29-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009196-0006-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 173:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6037323e
US-08-940-093-173

Query Match 30.2%; Score 29; DB 3; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NLADRMGL 11
|||:|
Db 8 NLLDRLLDL 16

RESULT 9
US-08-940-096-173
; Sequence 173, Application US/08940096
; Patent No. 6046166
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

Query Match 30.2%; Score 29; DB 3; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NLADRMGL 11
|||:|
Db 8 NLLDRLLDL 16

RESULT 10
US-09-465-719-173
; Sequence 173, Application US/09465719
; Patent No. 6265377
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940.096
FILING DATE: 29-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0005-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 173:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6046166e
US-08-940-096-173

Query Match 30.2%; Score 29; DB 3; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NLADRMGL 11
|||:|
Db 8 NLLDRLLDL 16

RESULT 10
US-09-465-719-173
; Sequence 173, Application US/09465719
; Patent No. 6265377
; GENERAL INFORMATION:
; APPLICANT: Dasseux, Jean-Louis
; APPLICANT: Sekul, Renate
; APPLICANT: Buttner, Klaus
; APPLICANT: Cornut, Isabelle
; APPLICANT: Metz, Gunther
; TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
; TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
; NUMBER OF SEQUENCES: 258
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/465,719
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/940,093
FILING DATE: 29-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0006-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 173:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6265377e
US-09-465-719-173

Query Match 30.2%; Score 29; DB 3; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NLADRMGL 11
DB 8 NLLDRLLD 16

RESULT 11

US-09-453-605-173
Sequence 173, Application US/09453605
Patent No. 6329341
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
Sekul, Renate
Buttner, Klaus
Cornut, Isabelle
Metz, Gunther
Dufourcq, Jean
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,605
FILING DATE: 26-No. 6329341-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/940,095
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0004-999
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 173:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6329341e
SEQUENCE DESCRIPTION: SEQ ID NO: 173:
US-09-453-605-173

Query Match 30.2%; Score 29; DB 4; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NLADRMGL 11
DB 8 NLLDRLLD 16

RESULT 12

US-09-453-838-173
Sequence 173, Application US/09453838
Patent No. 6376464
GENERAL INFORMATION:
APPLICANT: Dasseux, Jean-Louis
Sekul, Renate
Buttner, Klaus
Cornut, Isabelle
Metz, Gunther
Dufourcq, Jean
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,838
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/940,095
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0004-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 173:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6376464e
US-09-453-838-173

Query Match 30.2%; Score 29; DB 4; Length 22;

Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NLADRMGL 11
|||:|:
Db 8 NLLDRLLDL 16

RESULT 13

US-08-940-136-173
; Sequence 173, Application US/08940136
; Patent No. 6518412

GENERAL INFORMATION:

APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Buttner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
APPLICANT: Dufourcq, Jean
TITLE OF INVENTION: GENE THERAPY APPROACHES TO
SUPPLY APOLIPOPROTEIN A-I AGONISTS AND THEIR
USE TO TREAT DYSLIPIDEMIC DISORDERS.
NUMBER OF SEQUENCES: 274

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY

COUNTRY: USA

ZIP: 10036-2811

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/940,136
FILING DATE: 29-SEP-1997

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 009196-0007-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 173:

SEQUENCE CHARACTERISTICS:

LENGTH: 22 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6518412e

US-08-940-136-173

Query Match 30.2%; Score 29; DB 4; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NLADRMGL 11
|||:|:
Db 8 NLLDRLLDL 16

RESULT 14

US-09-453-841-173

; Sequence 173, Application US/09453841

; Patent No. 6573239

GENERAL INFORMATION:

APPLICANT: Dasseux, Jean-Louis
APPLICANT: Sekul, Renate
APPLICANT: Buttner, Klaus
APPLICANT: Cornut, Isabelle
APPLICANT: Metz, Gunther
APPLICANT: Dufourcq, Jean
TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
NUMBER OF SEQUENCES: 258

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY

COUNTRY: USA

ZIP: 10036-2811

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,841

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/940,095

FILING DATE: 29-SEP-1997

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 009196-0004-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-493-4935

TELEFAX: 650-493-5556

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 173:

SEQUENCE CHARACTERISTICS:

LENGTH: 22 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6573239e

US-09-453-841-173

Query Match 30.2%; Score 29; DB 4; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NLADRMGL 11
|||:|:
Db 8 NLLDRLLDL 16

RESULT 15

US-09-453-833-173

; Sequence 173, Application US/09453833

; Patent No. 6602854

GENERAL INFORMATION:

APPLICANT: Dasseux, Jean-Louis

APPLICANT: Sekul, Renate

APPLICANT: Buttner, Klaus

APPLICANT: Cornut, Isabelle

APPLICANT: Metz, Gunther

APPLICANT: Dufourcq, Jean

TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS

AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS

NUMBER OF SEQUENCES: 258

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

? COUNTRY: USA
? ZIP: 10036-2811
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FastSEQ Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/453,833
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/940,095
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Coruzzi, Laura A
? REGISTRATION NUMBER: 30,742
? REFERENCE/DOCKET NUMBER: 009196-0004-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650-493-4935
? TELEFAX: 650-493-5556
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 173:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 22 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: No. 6602854e
US-09-453-833-173

Query Match 30.2%; Score 29; DB 4; Length 22;
Best Local Similarity 66.7%; Pred.No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NLADRLGL 11
Db 8 NLLDRLLDL 16

Search completed: July 15, 2004, 15:13:22
Job time : 13.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:11:48 ; Search time 32.5333 Seconds
(without alignments)
192.148 Million cell updates/sec

Title: US-09-171-432A-46
Perfect score: 96
Sequence: 1 KINLADMLGLSGVQEIKEQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 288454

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	20	10	US-09-171-432A-46
2	96	100.0	25	10	US-09-171-432A-48
3	34	35.4	30	16	US-10-437-963-186323
4	33	34.4	27	14	US-10-195-730-291
5	32	33.3	20	14	US-10-112-488-53
6	32	33.3	20	16	US-10-673-860-40
7	32	33.3	24	14	US-10-271-078-12
8	32	33.3	25	16	US-10-416-249-149
9	31	32.3	20	14	US-10-087-464-3
10	31	32.3	24	15	US-10-424-233-44
11	30	31.2	12	14	US-10-057-789-189
12	30	31.2	12	14	US-10-212-628-189
13	30	31.2	20	9	US-09-731-221-71
14	30	31.2	21	14	US-10-206-155-1
15	30	31.2	21	14	US-10-360-836-3
					Sequence 46, Appl
					Sequence 48, Appl
					Sequence 186323,
					Sequence 291, Appl
					Sequence 53, Appl
					Sequence 40, Appl
					Sequence 12, Appl
					Sequence 149, Appl
					Sequence 3, Appl
					Sequence 44, Appl
					Sequence 189, Appl
					Sequence 189, Appl
					Sequence 71, Appl
					Sequence 1, Appl
					Sequence 3, Appl

16	30	31.2	25	16	US-10-416-249-151	Sequence 151, App
17	30	31.2	25	16	US-10-416-249-152	Sequence 152, App
18	30	31.2	29	12	US-10-424-599-146881	Sequence 146881,
19	29	30.2	17	14	US-10-057-789-286	Sequence 286, App
20	29	30.2	17	14	US-10-212-628-286	Sequence 286, App
21	29	30.2	20	10	US-09-171-432A-24	Sequence 24, Appl
22	29	30.2	22	10	US-09-865-989-173	Sequence 173, App
23	29	30.2	22	12	US-09-865-989-173	Sequence 173, App
24	29	30.2	22	14	US-10-099-574A-173	Sequence 173, App
25	29	30.2	22	14	US-10-147-849-173	Sequence 173, App
26	29	30.2	22	15	US-10-099-836B-173	Sequence 173, App
27	29	30.2	22	15	US-10-283-599-173	Sequence 173, App
28	29	30.2	25	16	US-10-416-249-153	Sequence 153, App
29	28	29.2	20	10	US-09-974-879-610	Sequence 610, App
30	28	29.2	20	10	US-09-305-736-611	Sequence 611, App
31	28	29.2	20	11	US-09-818-683-611	Sequence 611, App
32	28	29.2	20	12	US-10-621-401-610	Sequence 610, App
33	28	29.2	20	14	US-10-169-297-38	Sequence 38, Appl
34	28	29.2	23	9	US-09-975-553-19	Sequence 19, Appl
35	28	29.2	23	14	US-10-125-792-24	Sequence 24, Appl
36	28	29.2	23	14	US-10-125-778-24	Sequence 24, Appl
37	28	29.2	23	14	US-10-270-795-19	Sequence 19, Appl
38	28	29.2	23	14	US-10-270-876-19	Sequence 19, Appl
39	28	29.2	23	14	US-10-268-051-9	Sequence 9, Appl
40	28	29.2	23	14	US-10-125-772-24	Sequence 24, Appl
41	28	29.2	23	15	US-10-411-076-24	Sequence 24, Appl
42	28	29.2	23	15	US-10-410-872-19	Sequence 19, Appl
43	28	29.2	23	15	US-10-410-885-26	Sequence 26, Appl
44	28	29.2	28	10	US-09-798-889-121	Sequence 121, App
45	28	29.2	28	10	US-09-776-724A-565	Sequence 265, App

ALIGNMENTS

RESULT 1

US-09-171-432A-46
; Sequence 46, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1-20
OTHER INFORMATION: /label= YK-1331
US-09-171-432A-46

Query Match 100.0%; Score 96; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRMGLSGVQEIKEQ 20
Db 1 KINLADRMGLSGVQEIKEQ 20

RESULT 2

US-09-171-432A-48
; Sequence 48, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudiyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..25
OTHER INFORMATION: /label= YK-1757
US-09-171-432A-48

Query Match 100.0%; Score 96; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRMGLSGVQEIKEQ 20
Db 6 KINLADRMGLSGVQEIKEQ 25

RESULT 3

US-10-437-963-186323
; Sequence 186323, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 186323
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_83133C.1.pep
US-10-437-963-186323

Query Match 35.4%; Score 34; DB 16; Length 30;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 NLADRMGLSGV 14
Db 3 NCYDRYVGLTGI 14

RESULT 4

US-10-195-730-291
; Sequence 291, Application US/10195730
; Publication No. US20030144492A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et. al
; TITLE OF INVENTION: 101 Human Secreted Proteins
; FILE REFERENCE: PZ017P1
; CURRENT APPLICATION NUMBER: US/10/195,730
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/281,976
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/060,837
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 390
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 291
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-730-291

Query Match 34.4%; Score 33; DB 14; Length 27;
Best Local Similarity 56.2%; Pred. No. 1.7e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 INLADRMGLSGVQEI 17
Db 2 IHLEKRSGLSETQII 17

```

; TITLE OF INVENTION: NOVEL HUMAN LEUCINE-RICH REPEAT DOMAIN CONTAINING PROTEIN, HLLRCR-
; FILE REFERENCE: D0157 NP
; CURRENT APPLICATION NUMBER: US/10/271,078
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: U.S. 60/328,478
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-271-078-12

Query Match      33.3%; Score 32; DB 14; Length 24;
Best Local Similarity 35.3%; Pred. No. 2.2e+02;
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY      3 NLADRMGLSGVQEIKE 19
Db      1 NLALLSLRNGIEDVQE 17

RESULT 8
US-10-416-249-149
; Sequence 149, Application US/10416249
; Publication No. US20040132121A1
; GENERAL INFORMATION:
; APPLICANT: Dalrymple, Brian P.
; APPLICANT: Kongsuwan, Kritaya
; APPLICANT: Wijffels, Gene L.
; APPLICANT: Jennings, Philip A.
; APPLICANT: Kemp, Gregory W.
; TITLE OF INVENTION: METHOD OF IDENTIFYING ANTIBACTERIAL
; FILE REFERENCE: CULLM42.001APC
; CURRENT APPLICATION NUMBER: US/10/416,249
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: PCT/AU01/01436
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: AU PR 1320
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: AU PR 2919
; PRIOR FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 678
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-416-249-149

Query Match      33.3%; Score 32; DB 16; Length 25;
Best Local Similarity 33.3%; Pred. No. 2.3e+02;
Matches 4; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY      9 LGLSCVQEIKEQ 20
Db      1 VGFGSLSEVRQE 12

RESULT 9
US-10-087-464-3
; Sequence 3, Application US/10087464
; Publication No. US20030059436A1
; GENERAL INFORMATION:
; APPLICANT: Chishti, Athar
; APPLICANT: Oh, Steven
; APPLICANT: Liu, David
; APPLICANT: Goel, Vikas
; APPLICANT: Li, Xuerong
; TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses Thereof
; FILE REFERENCE: S1237/7019
; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: PROCESS FOR PRODUCING TRANSLUTAMINASE
; FILE REFERENCE: 219286USOCONT
; CURRENT APPLICATION NUMBER: US/10/112,488
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: PCT/JP00/06780
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: JP2000-280098
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: JP11-280098
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptovorticillium mobaraense
US-10-112-488-53

Query Match      33.3%; Score 32; DB 14; Length 20;
Best Local Similarity 27.8%; Pred. No. 1.8e+02;
Matches 5; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY      3 NLADRMGLSGVQEIKEQ 20
Db      3 DIKDRILKIPGMKFVEEK 20

RESULT 6
US-10-673-860-40
; Sequence 40, Application US/10673860
; Publication No. US20040126847A1
; GENERAL INFORMATION:
; APPLICANT: Ajinomoto Co., Inc.
; FILE REFERENCE: Y1J0182
; CURRENT APPLICATION NUMBER: US/10/673,860
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: JP 2001-98808
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptovorticillium mobaraense
US-10-673-860-40

Query Match      33.3%; Score 32; DB 16; Length 20;
Best Local Similarity 27.8%; Pred. No. 1.8e+02;
Matches 5; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY      3 NLADRMGLSGVQEIKEQ 20
Db      3 DIKDRILKIPGMKFVEEK 20

RESULT 7
US-10-271-078-12
; Sequence 12, Application US/10271078
; Publication No. US20030186267A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
```

; CURRENT APPLICATION NUMBER: US/10/087,464
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 06/272,930
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-464-3

Query Match 32.3%; Score 31; DB 14; Length 20;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 14 VQEIKEQ 20
:|:|:|:
Db 12 IQEVKEQ 18

RESULT 10
US-10-424-233-44
; Sequence 44, Application US/10424233
; Publication No. US20030220263A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN LEUCINE-RICH REPEAT-CONTAINING PROTEINS SPECIFICALLY
; TITLE OF INVENTION: EXPRESSED IN THE NERVOUS SYSTEM
; FILE REFERENCE: D0233 NP
; CURRENT APPLICATION NUMBER: US/10/424,233
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: U.S. 60/375,335
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-424-233-44

Query Match 32.3%; Score 31; DB 15; Length 24;
Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 9 LGLSGVQEIK 18
||:|:|:
Db 7 LGNNGLEIR 16

RESULT 11
US-10-057-789-189
; Sequence 189, Application US/10057789
; Publication No. US20030082522A1
; GENERAL INFORMATION:
; APPLICANT: Paul Haynes
; APPLICANT: Jing Wei
; APPLICANT: John Yates
; APPLICANT: Nancy Andon
; TITLE OF INVENTION: DIFFERENTIAL LABELING FOR QUANTITATIVE
; TITLE OF INVENTION: ANALYSIS OF COMPLEX PROTEIN MIXTURES
; FILE REFERENCE: NADII.022A
; CURRENT APPLICATION NUMBER: US/10/057,789
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 60/264,576
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/305,232
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189
; LENGTH: 12

; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Xaa = Modified Cysteine
US-10-057-789-189

Query Match 31.2%; Score 30; DB 14; Length 12;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 9 LGLSGVQEIK 18
||:|:|:
Db 3 LGITGVDQVR 12

RESULT 12
US-10-212-628-189
; Sequence 189, Application US/10212628
; Publication No. US20030087329A1
; GENERAL INFORMATION:
; APPLICANT: Paul Haynes
; APPLICANT: Jing Wei
; APPLICANT: John Yates
; APPLICANT: Nancy Andon
; TITLE OF INVENTION: DIFFERENTIAL LABELING FOR QUANTITATIVE
; TITLE OF INVENTION: ANALYSIS OF COMPLEX PROTEIN MIXTURES
; FILE REFERENCE: NADII.022CP1
; CURRENT APPLICATION NUMBER: US/10/212,628
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/264,576
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/305,232
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 10/057,789
; PRIOR FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 189
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Xaa = Modified Cysteine
US-10-212-628-189

Query Match 31.2%; Score 30; DB 14; Length 12;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 9 LGLSGVQEIK 18
||:|:|:
Db 3 LGITGVDQVR 12

RESULT 13
US-09-731-221-71
; Sequence 71, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 20
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-71

Query Match 31.2%; Score 30; DB 9; Length 20;
Best Local Similarity 33.3%; Pred. No. 3.9e+02;
Matches 6; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 3 NLADRMGLGSLGVQVEIK 20
|:|:|:|:|:|:|:|:|:|:
Db 2 NQSDRLQGRQEQQPFKE 19

RESULT 14
US-10-206-155-1
; Sequence 1, Application US/10206155
; Publication No. US20030157135A1
; GENERAL INFORMATION:
; APPLICANT: Tsuji, Moriya
; APPLICANT: Gonzalez-Asequinolaza, Gloria
; APPLICANT: Nussenzweig, Ruth S.
; APPLICANT: Kozuka, Yasuhiko
; TITLE OF INVENTION: USE OF GLYCOSYLCERAMIDES AS ADJUVANTS
; TITLE OF INVENTION: FOR VACCINES AGAINST INFECTIONS AND CANCER
; FILE REFERENCE: 5986/1H9S8US1
; CURRENT APPLICATION NUMBER: US/10/206,155
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/308,056
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 21
; TYPE: PRT
; ORGANISM: P. yoelii
US-10-206-155-1

Query Match 31.2%; Score 30; DB 14; Length 21;
Best Local Similarity 44.4%; Pred. No. 4.2e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Qy 3 NLADRMGLG--LSGVQVEIK 18
|:|:|:|:|:|:|:|:|:|:
Db 4 NIVNRLGLDALNGKPEEK 21

RESULT 15
US-10-360-836-3
; Sequence 3, Application US/10360836
; Publication No. US20030185854A1
; GENERAL INFORMATION:
; APPLICANT: Zavala, Fidel
; APPLICANT: Birkett, Ashlev
; TITLE OF INVENTION: USE OF RECOMBINANT HEPATITIS B CORE
; TITLE OF INVENTION: PARTICLES TO DEVELOP VACCINES AGAINST INFECTIOUS PATHOGENS
; TITLE OF INVENTION: AND MALIGNANCIES
; FILE REFERENCE: 5986/1J876
; CURRENT APPLICATION NUMBER: US/10/360,836
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/354,963
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Plasmodium
US-10-360-836-3

Query Match 31.2%; Score 30; DB 14; Length 21;
Best Local Similarity 44.4%; Pred. No. 4.2e+02;

Matches 8; Conservative 4; Mismatches 4; Indels 2; Gaps 1;
Qy 3 NLADRMGLG--LSGVQVEIK 18
|:|:|:|:|:|:|:|:|:|:
Db 4 NIVNRLGLDALNGKPEEK 21

Search completed: July 15, 2004, 15:22:58
Job time : 33.533 secs

117

THIS PAGE BLANK (USPTO)

117

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:36:31 ; Search time 50.1667 Seconds
(without alignments)
140.804 Million cell updates/sec

Title: US-09-171-432A-47
Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPPRKMGKLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 581562

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseqp_29Jan04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	129	100.0	25	2	AAW42930	Immunogen
2	129	100.0	26	4	AAB69447	Synthetic
3	103	79.8	20	2	AAW42924	Immunogen
4	103	79.8	21	4	AAB69441	Synthetic
5	65	50.4	20	2	AAW42925	Immunogen
6	65	50.4	21	4	AAB69442	Synthetic
7	46.5	36.0	23	2	AAR15517	BNP-8 pep
8	45	34.9	20	2	AAW42923	Immunogen
9	45	34.9	21	4	AAB69440	Synthetic
10	38	29.5	30	5	AAU84430	HIV GAG s
11	37	28.7	16	7	ADD71531	HLA-DP4 b
12	37	28.7	21	6	AAE31752	Androgen
13	37	28.7	30	4	AAB61644	MART 1 pe
14	37	28.7	30	5	AAU84871	Human MAR
15	36	27.9	19	2	AAW65529	Hepatitis
16	36	27.9	27	4	ABBS0903	Human sec
17	36	27.9	27	6	ABO45160	Novel hum
18	36	27.9	27	7	ABO26640	Protein a
19	36	27.9	30	4	ABBA1296	Peptide #
20	36	27.9	30	4	AAAM35081	Peptide #
21	36	27.9	30	4	ABB25267	Protein #
22	36	27.9	30	4	AAW74965	Human bra
23	36	27.9	30	4	AAW62161	Human bra
24	36	27.9	30	4	ABG56738	Human liv
25	36	27.9	30	5	ABG44704	Human pep

RESULT 1
AAW42930
ID AAW42930 standard; peptide; 25 AA.

AC AAW42930;
XX
DT 28-APR-1998 (first entry)
XX

DE Immunogenic Hepatitis A virus peptide YK-1665.
XX

KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.
XX

OS Synthetic.
OS Hepatitis A virus.
XX

PN WO9740147-A1.
XX

PD 30-OCT-1997.
XX

PF 18-APR-1997; 97WO-US006891.
XX

PR 19-APR-1996; 96US-0015644P.
XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX

PI Fields HA, Khudyakov YE;
XX

DR WPI; 1997-535831/49.
XX

PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
against HAV in a mammal.
XX

PS Claim 18; Page 112; 140pp; English.
XX

CC Peptides AAW42922-30 are immunogenic peptides corresponding to
immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
substantially similar to a portion of the amino acid sequence of the P2A
protein of HAV corresponding to amino acids 792-980. Compositions
containing the peptides can be used to induce an immune response to HAV
in a mammal. The peptides can also be used to detect the presence of
antibodies against HAV in mammalian serum. The peptides can also be used
to make an antibody against HAV by administering the peptide to a mammal
XX

SQ Sequence 25 AA;
XX

Query Match 100.0%; Score 129; DB 2; Length 25;

```

Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

RESULT 2
AAB69447
ID AAB69447 standard; peptide; 26 AA.
XX AC AAB69447;
XX AC
DT 20-APR-2001 (first entry)
XX
XX Synthetic HAV P2A peptide, SEQ ID NO: 47.
XX
XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
XX antigen; major structural capsid polypeptide; HAV antibody detection.
XX
XX Hepatitis A virus.
XX Synthetic.
XX
XX WO200105824-A2.
XX
XX WO200105824-A2.
XX
XX 25-JAN-2001.
XX
XX 14-JUL-2000; 2000WO-US019267.
XX
XX 15-JUL-1999; 99US-0144412P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Fields HA, Khudyakov YE;
XX
XX WPI, 2001-112681/12.
XX
XX Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines.
XX
XX Claim 13; Page 98; 130pp; English.
XX
XX The present sequence is one of a number of synthetic peptides which are
XX immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
XX comprise antigenic epitopes of the major structural capsid polypeptides
XX or non-structural polypeptides of HAV with one or more glutamine
XX molecules at the carboxy end of the peptide. The peptides are used to
XX detect the presence of antibodies against HAV in mammalian serum, to
XX detect the presence of HAV in a human or animal through the binding of
XX the peptide to an antibody, to detect acute phase infection by detecting
XX IgM antibodies in mammalian serum and detecting convalescence in a
XX mammal. The peptides are used to detect or quantify HAV antibodies in
XX samples in clinical or research-based assays using immunoblotting,
XX fluorescence in situ hybridisation analysis, gel-mobility shift assays,
XX tracking of radioactive or bioluminescent markers, chromatography or
XX electrophoresis. The peptides are used to induce an immune response to
XX HAV when administered to a human or animal. Glutamine at the carboxy end
XX of the peptides enhances the IgM antibody reactivity
XX
XX Sequence 26 AA;

Query Match 100.0%; Score 129; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRKMKGLF 25
Db 1 QRLKYAQEELSNEVLPPPRKMKGLF 25

RESULT 3
AAM42924

```

ID	AAW42924 standard; peptide; 20 AA.
XX	
XX	AAW42924;
XX	
DT	28-APR-1998 (first entry)
XX	
DE	Immunogenic Hepatitis A virus peptide YK-1317.
XX	
KW	Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW	antibody.
XX	
OS	Synthetic.
OS	Hepatitis A virus.
XX	
PN	WO9740147-A1.
XX	
PD	30-OCT-1997.
XX	
PF	18-APR-1997; 97WO-US006891.
XX	
PR	19-APR-1996; 96US-0015644P.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Fields HA, Khudyakov YE;
XX	
DR	WPI; 1997-535831/49.
XX	
PT	Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT	response to HAV in a mammal or to detect the presence of antibodies
PT	against HAV in a mammal.
XX	
PS	Claim 18; Page 112; 140pp; English.
XX	
CC	Peptides AAW42922-30 are immunogenic peptides corresponding to
CC	immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC	substantially similar to a portion of the amino acid sequence of the P2A
CC	protein of HAV corresponding to amino acids 792-980. The present peptide
CC	is derived from amino acids 810-829, and has a reactivity of 83.3% with
CC	acute sera. Compositions containing the peptides can be used to induce an
CC	immune response to HAV in a mammal. The peptides can also be used to
CC	detect the presence of antibodies against HAV in mammalian serum. The
CC	peptides can also be used to make an antibody against HAV by
CC	administering the peptide to a mammal
XX	
SQ	Sequence 20 AA;
	Query Match 79.8%; Score 103; DB 2; Length 20;
	Best Local Similarity 100.0%; Pred.No. 1e-08;
	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 QRLKYAQEELSNEVLPPPRK 20
Dd	1 QRLKYAQEELSNEVLPPPRK 20
RESULT 4	
AAB69441	ID
XX	AAB69441 standard; peptide; 21 AA.
XX	
AC	AAB69441;
XX	
DT	20-APR-2001 (first entry)
XX	
DE	Synthetic HAV P2A peptide, SEQ ID NO: 41.
XX	
KW	Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW	antigen; major structural capsid polypeptide; HAV antibody detection.
XX	
OS	Hepatitis A virus.
OS	Synthetic.
XX	
PN	WO200105824-A2.

XX PD 25-JAN-2001.
XX PF 14-JUL-2000; 2000WO-US019267.
XX PR 15-JUL-1999; 99US-0144412P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;
XX DR WPI; 2001-112681/12.
XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.
XX Claim 13; Page 94; 130pp; English.
XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the IGM antibody reactivity
XX Sequence 21 AA;
Query Match 79.8%; Score 103; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 1 QRLKYAQEELSNEVLPPPRK 20
RESULT 5
AAW42925
ID AAW42925 standard; peptide; 20 AA.
XX AC AAW42925;
XX DT 28-APR-1998 (first entry)
XX DE Immunogenic Hepatitis A virus peptide YK-1318.
XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.
XX OS Synthetic.
XX OS Hepatitis A virus.
XX PN WO9740147-A1.
XX PD 30-OCT-1997.
XX PF 18-APR-1997; 97WO-US006891.
XX PR 19-APR-1996; 96US-0015644P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;

XX WPI; 1997-535831/49.
XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.
XX Claim 18; Page 112; 140pp; English.
XX Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 823-942, and has a reactivity of 31.3% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal
XX Sequence 20 AA;
Query Match 50.4%; Score 65; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 14 VLPPPRKMKGLF 25
DB 1 VLPPPRKMKGLF 12
RESULT 6
AAB69442
ID AAB69442 standard; peptide; 21 AA.
XX AC AAB69442;
XX DT 20-APR-2001 (first entry)
XX DE Synthetic HAV P2A peptide, SEQ ID NO: 42.
XX KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
XX antigen; major structural capsid polypeptide; HAV antibody detection.
XX OS Hepatitis A virus.
XX OS Synthetic.
XX PN WO200105824-A2.
XX PD 25-JAN-2001.
XX PF 14-JUL-2000; 2000WO-US019267.
XX PR 15-JUL-1999; 99US-0144412P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;
XX DR WPI; 2001-112681/12.
XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.
XX Claim 13; Page 95; 130pp; English.
XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the IGM antibody reactivity
XX Sequence 21 AA;
Query Match 79.8%; Score 103; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 1 QRLKYAQEELSNEVLPPPRK 20
RESULT 5
AAW42925
ID AAW42925 standard; peptide; 20 AA.
XX AC AAW42925;
XX DT 28-APR-1998 (first entry)
XX DE Immunogenic Hepatitis A virus peptide YK-1318.
XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.
XX OS Synthetic.
XX OS Hepatitis A virus.
XX PN WO9740147-A1.
XX PD 30-OCT-1997.
XX PF 18-APR-1997; 97WO-US006891.
XX PR 19-APR-1996; 96US-0015644P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;

CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IGM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the Igm antibody reactivity
 XX Sequence 21 AA;

SQ Query Match 50.4%; Score 65; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.0078;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 VLPPPRKMKGLF 25
 Db |||||

RESULT 7
 AAR15517
 ID AAR15517 standard; protein; 23 AA.

XX AAR15517;

XX 25-MAR-2003 (revised)
 DT 09-MAR-1992 (first entry)

XX BMP-8 peptide.

XX Cartilage; wound healing; tissue repair; BMP.

XX Bos taurus.

XX WO9118098-A.

XX 28-NOV-1991.

XX 16-MAY-1990; 90US-00525357.

XX 16-MAY-1990; 90US-00525357.

XX 15-JAN-1991; 91US-00641204.

XX (GEMY) GENETICS INST INC.

XX Hewick RM, Wang JH;

XX WPI; 1991-369252/50.

XX N-PSDB; AAQ15240, AAQ15242, AAQ15245, AAQ15246.

XX New BMP-8 protein - useful in inducing cartilage and/or bone formation to
 PT treat wounds and repair fractures and tissues, e.g. burns, incisions and
 PT ulcers.

XX Claim 1(d); Page 46; 50pp; English.

XX This sequence shares some homology (i.e. Aen-Glu-Leu-Pro) with BMP-3 (see
 CC WO88/00205 and WO89/10409). Pharmaceutical compsns. contg. BMP-8, which
 CC comprises at least one of the fragments represented in AAR15517 and
 CC AAR15522, can be used to aid bone and/or cartilage formation or wound
 CC healing and tissue repair. The proteins are not very species specific so
 CC can be used in domestic and farm animals as well as humans. See also
 CC AAQ15240-48, AAR15517 and AAR15522. (Updated on 25-MAR-2003 to correct PA
 CC field.)

XX Sequence 23 AA;

XX Query Match 36.0%; Score 46.5; DB 2; Length 23;
 XX Best Local Similarity 60.0%; Pred. No. 6;
 XX Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 11 SNEVLPPPRKMKGLF 25
 Db 1 TNE-LPPPNKLPGLF 14
 |||||

RESULT 8
 AAW42923
 ID AAW42923 standard; peptide; 20 AA.

XX AAW42923;

XX 28-APR-1998 (first entry)

XX Immunogenic Hepatitis A virus peptide YK-1316.

XX Immunogenic peptide; immunogenic epitope; P2A protein; immune response;

XX antibody.

XX Synthetic.

XX Hepatitis A virus.

XX WO9740147-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-US006891.

XX 19-APR-1996; 96US-0015644P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 1997-535831/49.

XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.

XX Claim 18; Page 112; 140pp; English.

XX Peptides AAW42922-30 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the P2A
 CC protein of HAV corresponding to amino acids 792-980. The present peptide
 CC is derived from amino acids 799-818, and has a reactivity of 41.7% with
 CC acute sera. Compositions containing the peptides can be used to induce an
 CC immune response to HAV in a mammal. The peptides can also be used to
 CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal

XX Sequence 20 AA;

XX Query Match 34.9%; Score 45; DB 2; Length 20;

XX Best Local Similarity 100.0%; Pred. No. 8.9;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEE 9
 |||||

Db 12 QRLKYAQEE 20
 |||||

RESULT 9

XX AAB69440 standard; peptide; 21 AA.

XX AAB69440;

XX 20-APR-2001 (first entry)

XX Synthetic HAV P2A peptide, SEQ ID NO: 40.

KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 XX
 OS Hepatitis A virus.
 OS Synthetic.
 XX
 PN WO200105824-A2.
 XX
 PD 25-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-US019267.
 XX
 PR 15-JUL-1999; 99US-0144412P.
 XX
 XX (USHS) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Fields HA, Khudyakov YE;
 XX
 DR WPI; 2001-112681/12.
 XX
 PT Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 XX
 PS Claim 13; Page 93; 130pp; English.
 XX
 CC The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IgM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IgM antibody reactivity
 XX
 SQ Sequence 21 AA;
 Query Match 34.9%; Score 45; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 9.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEE 9
 Db 12 QRLKYAQEE 20
 RESULT 10
 AAU84430
 ID AAU84430 standard; peptide; 30 AA.
 XX
 AC AAU84430;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE HIV GAG segment 32.
 XX
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
 KW viral infection; human immunodeficiency virus; melanoma;
 KW bacterial infection; Salmonella; Legionella; parasitic infection;
 KW Trypanosoma; Toxoplasma; Giardia.
 XX
 OS Human immunodeficiency virus 1.
 OS Synthetic.
 XX
 PN WO200190197-A1.
 XX

PD 29-NOV-2001.
 XX
 PF 25-MAY-2001; 2001WO-AU000622.
 XX
 PR 26-MAY-2000; 2000AU-00007761.
 XX
 PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Thomson SA, Ramshaw IA;
 XX
 DR WPI; 2002-147575/19.
 DR N-PSDB; ABK36269.
 XX
 PT New synthetic polypeptides having several different segments of at least
 PT one parent polypeptide linked together differently compared to the
 PT linkage in the parent polypeptide, for inducing immune response against a
 PT pathogen or cancer.
 XX
 PS Example 1; Fig 12; 364pp; English.
 XX
 CC The invention relates to a new synthetic polypeptide (1) comprising
 CC several different segments of at least one parent polypeptide linked
 CC together in a different relationship relative to their linkage in the
 CC parent polypeptide to impede, abrogate or otherwise alter at least one
 CC function associated with the parent polypeptide and for inducing an
 CC immune response against a pathogen or cancer. Also included are a
 CC synthetic polynucleotide encoding and a computer system for designing the
 CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
 CC are referred to as a Savine. The synthetic polypeptide is useful for
 CC modulating immune responses preferably directed against a pathogen or a
 CC cancer. (e.g., cancers of the lung, breast, ovary, cervix, colon, head
 CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
 CC oesophagus, brain, testicle, uterus), as potentiating agents.
 CC Compositions comprising the polypeptide may be used in the treatment or
 CC prophylaxis against viral (such as infections caused by HIV (human
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
 CC Salmonella, Streptococcal, Legionella and Mycobacterium) or parasitic
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
 CC a peptide derived from a parent protein used to construct a savine of the
 CC invention
 XX
 SQ Sequence 30 AA;
 Query Match 29.5%; Score 38; DB 5; Length 30;
 Best Local Similarity 42.1%; Pred. No. 1.6e+02;
 Matches 8; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
 QY 7 QEELSNEVLPPPRKMKGLF 25
 Db 9 QEXKDKEXYPPXASLKSLF 27
 RESULT 11
 ADD71531
 ID ADD71531 standard; peptide; 16 AA.
 XX
 AC ADD71531;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE HLA-DP4 binding peptide ligand #93.
 XX
 KW cytostatic; immunostimulant; immunosuppressive; neuroprotective;
 KW antidiabetic; antiallergic; ligand; HLA-DP4; human leukocyte antigen;
 KW immunomodulator; vaccine; pathogen; tumor cell; multiple sclerosis;
 KW diabetes; allergy; graft rejection.
 XX
 OS Synthetic.
 XX
 PN FR2830940-A1.

PT New antigenic compositions for eliciting improved immune responses,
 PT especially in overcoming tolerance to self-antigens.

XX Disclosure; Page 18; 40pp; English.

CC The present invention relates to a composition for eliciting an immune
 CC response to a target molecule comprising an immunological adjuvant and
 CC one or more peptides. The present sequence is one such peptide. The
 CC peptides are similar to, but different from, a target peptide e.g. MART
 CC 1. MART 1 is a protein associated with melanoma. The present peptide can
 CC elicit the production of antibodies against MART 1, and so can be used to
 CC treat melanoma

XX Sequence 30 AA;

Query Match 28.7%; Score 37; DB 4; Length 30;
 Best Local Similarity 61.5%; Pred. No. 2.3e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 AOEELSNVEVLPPP 18
 | : | | | | |
 Db 15 AYEKLSAEQSPPP 27

RESULT 14

AAU84871
 ID AAU84871 standard; peptide; 30 AA.

AC AAU84871;

DT 08-MAY-2002 (first entry)
 DE Human MART segment 7.

KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
 KW viral infection; human immunodeficiency virus; melanoma;
 KW bacterial infection; Salmonella; Legionella; parasitic infection;
 KW Trypanosoma; Toxoplasma; Giardia.

XX Homo sapiens.

OS WO200190197-A1.

PN 29-NOV-2001.

PD 25-MAY-2001; 2001WO-AU000622.

PF 26-MAY-2000; 2000AU-00007761.

XX (AUSU) UNIV AUSTRALIAN NAT.

PA Thomson SA, Ramshaw IA;

PI WPI; 2002-147575/19.

DR N-PSDB; ABK36691.

XX New synthetic polypeptides having several different segments of at least
 PT one parent polypeptide linked together differently compared to the
 PT linkage in the parent polypeptide, for inducing immune response against a
 PT pathogen or cancer.

PS Example 3; Fig 27; 364pp; English.

CC The invention relates to a new synthetic polypeptide (I) comprising
 CC several different segments of at least one parent polypeptide linked
 CC together in a different relationship relative to their linkage in the
 CC parent polypeptide to impede, abrogate or otherwise alter at least one
 CC function associated with the parent polypeptide and for inducing an
 CC immune response against a pathogen or cancer. Also included are a
 CC synthetic polynucleotide encoding and a computer system for designing the
 CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
 CC are referred to as a Savine. The synthetic polypeptide is useful for
 CC modulating immune responses preferably directed against a pathogen or a

CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
 CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
 CC oesophagus, brain, testicle, uterus), as potentiating agents.
 CC Compositions comprising the polypeptide may be used in the treatment or
 CC prophylaxis against viral (such as infections caused by HIV (human
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
 CC a peptide derived from a parent protein used to construct a savine of the
 CC invention

XX Sequence 30 AA;

Query Match 28.7%; Score 37; DB 5; Length 30;
 Best Local Similarity 61.5%; Pred. No. 2.3e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 AOEELSNVEVLPPP 18
 | : | | | | |
 Db 15 AYEKLSAEQSPPP 27

RESULT 15

AAW65529
 ID AAW65529 standard; peptide; 19 AA.

XX AC AAW65529;

DT 27-AUG-2003 (revised)

DT 12-OCT-1998 (first entry)

XX Hepatitis B surface antigen derived peptide linked to polylysine core.
 XX Annexin V; hepatitis B surface antigen; immunogen; vaccine;
 KW hepatitis delta virus; infection; HBsAg; polylysine core.

OS Synthetic.

OS Hepatitis B virus.

EH Key Location/Qualifiers
 FT 1. .13

FT /note= "peptide derived from hepatitis B surface antigen
 FT is attached to polylysine core via the alpha amino group
 FT of Lys(14); a second copy of the 13-mer is linked to
 FT Lys(14) via the omega amino group"

FT Modified-site 14

FT /note= "Lys(14) is linked to one copy of the peptide
 FT through the alpha amino group, and to a second copy of
 FT the peptide (not shown) via the omega amino group"

FT Modified-site 15

FT /note= "the alpha amino group of Lys(15) forms a peptide
 FT linkage with the carboxyl amino group of Lys(14); the
 FT omega amino group of Lys(15) forms a peptide bond with a
 FT second Lys residue analogous to Lys(14)"

FT Modified-site 17

FT /note= "bio-Lys"

XX WO9829442-A1.

XX 09-JUL-1998.

XX 23-DEC-1997; 97WO-EP007268.

XX 30-DEC-1996; 96EP-00870164.

PR 11-JUL-1997; 97EP-00870103.

XX (INNO-) INNOGENETICS NV.

XX Depla E, Maertens G, Yap S, De Meyer S;

DR WPI; 1998-388040/33.

XX Immunogenic polypeptide from hepatitis B surface antigen - useful in,
PT e.g. vaccine against hepatitis B virus or hepatitis delta virus
PT infection.

XX
XX
PS Example 5; Page 36; 71pp; English.

XX The invention relates to an immunogenic peptide derived from hepatitis B
CC surface antigen (HBsAg) which competes with the hepatitis B surface
CC antigen/annexin V interaction or which binds a compound or antibody
CC competing with the hepatitis B surface antigen/annexin V interaction.
CC Also claimed are: (1) a combination of the immunogenic peptide and a
CC negatively charged phospholipid; (2) a peptide composition comprising the
CC immunogenic peptide; (3) a vaccine comprising the immunogenic peptide as
CC an active substance; (4) antibodies which specifically bind to the
CC peptide and inhibit binding of HBsAg to annexin V, and (5) a therapeutic
CC composition comprising as an active substance the antibodies of (4). The
CC vaccine of (3), and the therapeutic composition of (5), can be used as an
CC inoculum to vaccinate humans against an infection with hepatitis B
CC and/or hepatitis delta virus. The immunogenic peptide can be used in a
CC method to detect antibodies which are capable of competing with the
CC hepatitis B and/or hepatitis delta virus surface antigen/annexin V
CC interaction. The immunogenic peptide can also be used to screen for drugs
CC which block the binding between annexin V and the peptide, and as a
CC therapeutic to treat humans infected with hepatitis B virus and/or
CC hepatitis delta virus. The present sequence represents a branched chain
CC polylysine of formula [(SVRVEQVRPPQK)2K]2-KGK(bio/GA). (Updated on 27-AUG
CC -2003 to correct OS field.)

XX Sequence 19 AA;

Query Match 27.9%; Score 36; DB 2; Length 19;
Best Local Similarity 54.5%; Pred. NO. 2e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 13 EVLPPPRKMKG 23
: : : : :
Db 6 QVVRPPQKKKG 16

Search completed: July 15, 2004, 14:46:52
Job time : 51.1667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:47:02 ; Search time 15.5 Seconds
(without alignments)
83.268 Million cell updates/sec

Title: US-09-171-432A-47
Perfect score: 129
Sequence: 1 QRLKYAELSNEVLPPRRKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 204689

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap: *
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap: *
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap: *
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap: *
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pap: *
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.5	36.0	23	US-07-800-364B-4	Sequence 4, Appli
2	46.5	36.0	23	PCT-US91-03388-4	Sequence 4, Appli
3	36	27.9	27	US-09-205-258-856	Sequence 856, App
4	35	27.1	22	US-09-822-624-11	Sequence 11, Appl
5	33	25.6	16	US-08-602-999A-199	Sequence 199, App
6	33	25.6	16	US-09-500-124-199	Sequence 199, App
7	33	25.6	20	US-08-817-548A-2	Sequence 2, Appli
8	33	25.6	23	US-08-268-251-52	Sequence 52, Appl
9	33	25.6	23	PCT-US93-01112-52	Sequence 52, Appl
10	33	25.6	25	US-09-540-224-8	Sequence 8, Appli
11	33	25.6	25	US-09-564-595D-40	Sequence 40, Appl
12	33	25.6	25	US-09-808-972-10	Sequence 10, Appl
13	33	25.6	27	US-08-262-037-107	Sequence 107, App
14	33	25.6	27	US-08-602-999A-339	Sequence 339, App
15	32.5	25.2	15	US-09-500-124-339	Sequence 339, App
16	32.5	25.2	15	US-08-444-818-589	Sequence 589, App
17	32	24.8	8	US-08-146-028-452	Sequence 452, App
18	32	24.8	9	US-08-146-028-452	Sequence 452, App
19	32	24.8	9	US-08-146-028-452	Sequence 452, App
20	32	24.8	9	US-08-723-425A-452	Sequence 452, App
21	32	24.8	9	US-08-723-425A-453	Sequence 453, App
22	32	24.8	9	US-09-112-206-452	Sequence 452, App
23	32	24.8	9	US-09-112-206-452	Sequence 452, App
24	32	24.8	9	US-09-730-497A-400	Sequence 400, App
25	32	24.8	9	US-09-730-497A-401	Sequence 401, App
26	32	24.8	9	US-09-790-497A-576	Sequence 576, App
27	32	24.8	9	US-09-576-824A-400	Sequence 400, App

28	32	24.8	9	4	US-09-576-824A-401	Sequence 401, App
29	32	24.8	11	4	US-09-576-824A-576	Sequence 576, App
30	32	24.8	13	2	US-08-572-951-13	Sequence 13, Appl
31	32	24.8	19	2	US-09-017-205-67	Sequence 67, Appl
32	32	24.8	20	2	US-08-466-975A-20	Sequence 20, Appl
33	32	24.8	20	2	US-08-391-671A-20	Sequence 20, Appl
34	32	24.8	20	3	US-08-467-902A-20	Sequence 20, Appl
35	32	24.8	20	3	US-09-275-265-20	Sequence 20, Appl
36	32	24.8	20	4	US-08-850-328-13	Sequence 13, Appl
37	32	24.8	20	4	US-09-941-611-20	Sequence 20, Appl
38	32	24.8	20	4	US-09-790-497A-63	Sequence 63, Appl
39	32	24.8	20	4	US-09-790-497A-103	Sequence 103, Appl
40	32	24.8	22	1	US-07-753-461A-8	Sequence 8, Appli
41	32	24.8	22	1	US-08-315-831A-8	Sequence 8, Appli
42	32	24.8	22	2	US-08-662-318-8	Sequence 8, Appli
43	32	24.8	22	2	US-08-146-028-63	Sequence 63, Appl
44	32	24.8	22	2	US-08-146-028-103	Sequence 103, App
45	32	24.8	22	3	US-08-723-425A-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-07-800-364B-4
; Sequence 4, Application US/07800364B
; Patent No. 5688678
; GENERAL INFORMATION:
; APPLICANT: Hewick, Rodney M.
; APPLICANT: Wang, Jack H.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/800,364B
; FILING DATE: 26-NOV-1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: GI 5182A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Bos taurus
; TISSUE TYPE: Bone
US-07-800-364B-4

Query Match 36.0% ; Score 46.5 ; DB 1 ; Length 23 ;
Best Local Similarity 60.0% ; Pred. No. 1 ;
Matches 9 ; Conservative 3 ; Mismatches 2 ; Indels 1 ; Gaps 1 ;

QY 11 SNEVLPPPRKMKGLF 25
Db 1 TNE-LPPPNKLPGLF 14

RESULT 2
PCT-US91-03388-4
Sequence 4, Application PC/TUS9103388
GENERAL INFORMATION:
APPLICANT: Hewick, Rodney M.
APPLICANT: Wang, Jack H.
TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/03388
FILING DATE: 19910515
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,245
REFERENCE/DOCKET NUMBER: G15182X-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-876-1170
TELEFAX: 617-876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Bone
PCT-US91-03388-4

Query Match 36.0%; Score 46.5; DB 5; Length 23;
Best Local Similarity 60.0%; Fred. No. 1;
Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 11 SNEVLPPPRKMKGLF 25
Db 1 TNE-LPPPNKLPGLF 14

RESULT 3
US-09-205-258-856
Sequence 856, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: PZ007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921

; EARLIER FILING DATE: 1998-07-15
 ; EARLIER APPLICATION NUMBER: 60/094,657
 ; EARLIER FILING DATE: 1998-07-30
 ; NUMBER OF SEQ ID NOS: 1227
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 856
 ; LENGTH: 27
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-205-258-856

Query Match 27.9%; Score 36; DB 4; Length 27;
 Best Local Similarity 60.0%; Pred. No. 53;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 15 LPPPRKMKGL 24
 DB 16 VPPPLKMPGV 25

RESULT 4
 US-09-822-624-11
 ; Sequence 11, Application US/09822624
 ; Patent No. 6440714
 ; GENERAL INFORMATION:
 ; APPLICANT: Abell, Creed W.
 ; APPLICANT: Kwan, Sua-Wah
 ; APPLICANT: Zhou, Binhua
 ; APPLICANT: Wo, Bo
 ; TITLE OF INVENTION: Tyr393 and Tyr398 Mutants of Monoamine Oxidase B
 ; FILE REFERENCE: D6237PCT
 ; CURRENT APPLICATION NUMBER: US/09/822,624
 ; CURRENT FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 2000-03-30
 ; PRIOR FILING DATE: US 60/193,178
 ; NUMBER OF SEQ ID NOS: 26
 ; SEQ ID NO 11
 ; LENGTH: 22
 ; TYPE: PRT
 ; ORGANISM: C. vinosium
 ; OTHER INFORMATION: Amino acid sequence flanking the covalent FAD
 US-09-822-624-11

Query Match 27.1%; Score 35; DB 4; Length 22;
 Best Local Similarity 46.7%; Pred. No. 61;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 10 LSNEVLPPPRKMKGL 24
 DB 6 LSNEVIGGDRKLES1 20

RESULT 5
 US-08-602-999A-199
 ; Sequence 199, Application US/08602999A
 ; Patent No. 6184205
 ; GENERAL INFORMATION:
 ; APPLICANT: SPARKS, Andrew B.
 ; APPLICANT: KAY, Brian K.
 ; APPLICANT: THORN, Judith M.
 ; APPLICANT: QUILLIAM, Lawrence A.
 ; APPLICANT: DER, Channing J.
 ; APPLICANT: FOWLKES, Dana M.
 ; APPLICANT: RIDER, James E.
 ; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
 ; TITLE OF INVENTION: ISOLATING AND USING SAME
 ; NUMBER OF SEQUENCES: 467
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.

; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/602,999A
 ; FILING DATE: 16-FEB-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mistrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-202
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 199:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 16 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 US-08-602-999A-199

Query Match 25.6%; Score 33; DB 3; Length 16;
 Best Local Similarity 75.0%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 16 PPRPKMKG 23
 DB 6 PPPPKMKG 13

RESULT 6
 US-09-500-124-199
 ; Sequence 199, Application US/09500124
 ; Patent No. 6432920
 ; GENERAL INFORMATION:
 ; APPLICANT: SPARKS, Andrew B.
 ; APPLICANT: KAY, Brian K.
 ; APPLICANT: THORN, Judith M.
 ; APPLICANT: QUILLIAM, Lawrence A.
 ; APPLICANT: DER, Channing J.
 ; APPLICANT: FOWLKES, Dana M.
 ; APPLICANT: RIDER, James E.
 ; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
 ; TITLE OF INVENTION: ISOLATING AND USING SAME
 ; NUMBER OF SEQUENCES: 467
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/500,124
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/602,999
 ; FILING DATE: 16-FEB-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mistrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-202

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 199:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-500-124-199

Query Match 25.6%; Score 33; DB 4; Length 16;
Best Local Similarity 75.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 16 PPRKMKG 23

Db 6 PPEIPMKG 13

RESULT 7
US-08-817-548A-2
Sequence 2, Application US/08817548A
Patent No. 6174532
GENERAL INFORMATION:
APPLICANT: MARIA SALLERIA CAMPO
APPLICANT: WILLIAM FLEMING HOGAN JARRETT
TITLE OF INVENTION: PAPILLONAVIRUS VACCINE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: ONE POST OFFICE SQUARE
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,548A
FILING DATE: 12 MAY 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MATTHEW P. VINCENT
REGISTRATION NUMBER: 36709
REFERENCE/DOCKET NUMBER: CFV-301.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
TELEX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-817-548A-2

Query Match 25.6%; Score 33; DB 3; Length 20;
Best Local Similarity 53.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 AOEELSNVLP 18

Db 2 AEETIAEVHP 14

RESULT 8
US-08-268-251-52

Sequence 52, Application US/08268251
Patent No. 5585475
GENERAL INFORMATION:
APPLICANT: Jamieson, Gordon A
APPLICANT: Dedman, John R
APPLICANT: Kaetzel, Marcia A
TITLE OF INVENTION: Calmodulin-Binding Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,251
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,219
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 272.001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 52:

SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-268-251-52

Query Match 25.6%; Score 33; DB 1; Length 23;
Best Local Similarity 44.4%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 10 LSNVLP 18

Db 12 MTNKAIPP 20

RESULT 9
PCT-US93-01112-52
Sequence 52, Application PC/TUS9301112
GENERAL INFORMATION:
APPLICANT: Jamieson, Gordon A
APPLICANT: Dedman, John R
APPLICANT: Kaetzel, Marcia A
TITLE OF INVENTION: Calmodulin-Binding Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01112
; FILING DATE: 19930208
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,219
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 272.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-01112-52

Query Match 25.6%; Score 33; DB 5; Length 23;
Best Local Similarity 44.4%; Pred. No. 1.3e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 10 LSNVLPPP 18
DB 12 MTNKAIPPP 20

RESULT 10
US-09-540-224-8
; Sequence 8, Application US/09540224
; Patent No. 6468543
; GENERAL INFORMATION:
; APPLICANT: Gilbertson, Debra G.
; TITLE OF INVENTION: METHODS FOR PROMOTING GROWTH OF BONE,
; TITLE OF INVENTION: LIGAMENT AND CARTILAGE USING ZVEGF4
; FILE REFERENCE: 00-28
; CURRENT APPLICATION NUMBER: US/09/540,224
; CURRENT FILING DATE: 2000-03-31
; EARLIER APPLICATION NUMBER: US 60/180,169
; EARLIER FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide
US-09-540-224-8

Query Match 25.6%; Score 33; DB 4; Length 25;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 QEELSNVLPPPP 19
DB 4 QEDLENNYLDTPR 16

RESULT 11
US-09-564-595D-40
; Sequence 40, Application US/09564595D
; Patent No. 6495668
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Teresa
; APPLICANT: Hart, Charles E.
; APPLICANT: Sheppard, Paul O.
```

```
; TITLE OF INVENTION: GROWTH FACTOR HOMOLOG ZVEGF4
; FILE REFERENCE: 99-19
; CURRENT APPLICATION NUMBER: US/09/564,595D
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: US 09/304,216
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: US 60/164,463
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: US 60/180,169
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide
US-09-564-595D-40

Query Match 25.6%; Score 33; DB 4; Length 25;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 QEELSNVLPPPP 19
DB 4 QEDLENNYLDTPR 16

RESULT 12
US-09-808-972-10
; Sequence 10, Application US/09808972
; Patent No. 6630142
; GENERAL INFORMATION:
; APPLICANT: Hart, Charles E.
; APPLICANT: Topouzis, Stavros
; APPLICANT: Gilbertson, Debra G.
; TITLE OF INVENTION: METHOD OF TREATING FIBROBLASTIC
; TITLE OF INVENTION: DISORDERS
; FILE REFERENCE: 00-79
; CURRENT APPLICATION NUMBER: US/09/808,972
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/235,295
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/564,595
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: US 60/180,169
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/164,463
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: US 60/132,250
; PRIOR FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide
US-09-808-972-10

Query Match 25.6%; Score 33; DB 4; Length 25;
Best Local Similarity 53.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 QEELSNVLPPPP 19
DB 4 QEDLENNYLDTPR 16

RESULT 13
US-08-262-037-107
```

```
; Sequence 107, Application US/08262037
; Patent No. 5747239
; GENERAL INFORMATION:
; APPLICANT: Chang Yi Wang and Barbara Hosein
; TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
; TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
; TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVE.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/262,037
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/719,819
; FILING DATE: 24-June-1991
; APPLICATION NUMBER: 07/667,275
; FILING DATE: 11-Mar-1991
; APPLICATION NUMBER: 07/651,735
; FILING DATE: 07-Feb-1991
; APPLICATION NUMBER: 07/558,799
; FILING DATE: 26-July-1990
; APPLICATION NUMBER: 07/510,153
; FILING DATE: 16-April-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C. H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4043 US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: Amino acid
; STRANDEDNESS:
; TOPOLOGY: Unknown
; US-08-262-037-107

Query Match 25.6%; Score 33; DB 1; Length 27;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 15 LPPPRKMK 22
DB 19 VPPPRKXR 26

RESULT 14
5498694-8
Patent No. 5498694
APPLICANT: RIUSLAHTI, ERKKI I.
TITLE OF INVENTION: PEPTIDES OF THE CYTOPLASMIC DOMAIN OF
INTEGRIN
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,967
FILING DATE: 10-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 973,547
FILING DATE: 09-NOV-1992
```

```
; APPLICATION NUMBER: 357,824
; FILING DATE: 25-MAY-1989
; SEQ ID NO:8;
; LENGTH: 27
; 5498694-8

Query Match 25.6%; Score 33; DB 6; Length 27;
Best Local Similarity 34.8%; Pred. No. 1.6e+02;
Matches 8; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMKG 23
DB 1 KVRPPQEQRQLEPHENGEG 23

RESULT 15
US-08-602-999A-339
; Sequence 339, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: OULLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James B.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 339:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-602-999A-339

Query Match 25.2%; Score 32.5; DB 3; Length 15;
Best Local Similarity 50.0%; Pred. No. 99;
Matches 8; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 7 QEELSNEVLPPPRKMK 22
DB 1 QEELPP---PPPOKKR 13

Search completed: July 15, 2004, 15:13:23
Job time : 16.5 secs
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:11:48 ; Search time 40.6667 Seconds
(without alignments)
192.148 Million cell updates/sec

Title: US-09-171-432a-47
Perfect score: 129
Sequence: 1 QRLKYAQEELSNEVLPPRRKMGFLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 288454

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	129	100.0	25	US-09-171-432A-47	Sequence 47, Appl
2	103	79.8	20	US-09-171-432A-41	Sequence 41, Appl
3	65	50.4	20	US-09-171-432A-42	Sequence 42, Appl
4	45	34.9	20	US-09-171-432A-40	Sequence 40, Appl
5	38	29.5	30	US-10-296-734-73	Sequence 73, Appl
6	37	28.7	21	US-10-097-175-3	Sequence 3, Appl
7	37	28.7	30	US-10-296-734-936	Sequence 936, Appl
8	36	27.9	27	US-09-933-767-856	Sequence 856, Appl
9	36	27.9	27	US-10-004-860-856	Sequence 856, Appl
10	36	27.9	27	US-10-023-282-856	Sequence 856, Appl
11	36	27.9	30	US-09-864-761-40565	Sequence 40565, A
12	35	27.1	13	US-10-226-007-832	Sequence 832, Appl
13	35	27.1	14	US-10-226-007-833	Sequence 833, Appl
14	35	27.1	14	US-10-226-007-840	Sequence 840, Appl
15	35	27.1	15	US-10-226-007-834	Sequence 834, Appl

16	35	27.1	15	14	US-10-226-007-841	Sequence 841, App
17	35	27.1	15	14	US-10-226-007-847	Sequence 847, App
18	35	27.1	16	14	US-10-226-007-835	Sequence 835, App
19	35	27.1	16	14	US-10-226-007-842	Sequence 842, App
20	35	27.1	16	14	US-10-226-007-848	Sequence 848, App
21	35	27.1	16	14	US-10-226-007-853	Sequence 853, App
22	35	27.1	17	14	US-10-226-007-836	Sequence 836, App
23	35	27.1	17	14	US-10-226-007-843	Sequence 843, App
24	35	27.1	17	14	US-10-226-007-849	Sequence 849, App
25	35	27.1	17	14	US-10-226-007-854	Sequence 854, App
26	35	27.1	17	14	US-10-226-007-858	Sequence 858, App
27	35	27.1	18	14	US-10-226-007-837	Sequence 837, App
28	35	27.1	18	14	US-10-226-007-844	Sequence 844, App
29	35	27.1	18	14	US-10-226-007-850	Sequence 850, App
30	35	27.1	18	14	US-10-226-007-855	Sequence 855, App
31	35	27.1	18	14	US-10-226-007-859	Sequence 859, App
32	35	27.1	18	14	US-10-226-007-862	Sequence 862, App
33	35	27.1	19	14	US-10-226-007-838	Sequence 838, App
34	35	27.1	19	14	US-10-226-007-845	Sequence 845, App
35	35	27.1	19	14	US-10-226-007-851	Sequence 851, App
36	35	27.1	19	14	US-10-226-007-856	Sequence 856, App
37	35	27.1	19	14	US-10-226-007-860	Sequence 860, App
38	35	27.1	19	14	US-10-226-007-863	Sequence 863, App
39	35	27.1	19	14	US-10-226-007-865	Sequence 865, App
40	35	27.1	20	14	US-10-226-007-839	Sequence 839, App
41	35	27.1	20	14	US-10-226-007-846	Sequence 846, App
42	35	27.1	20	14	US-10-226-007-852	Sequence 852, App
43	35	27.1	20	14	US-10-226-007-857	Sequence 857, App
44	35	27.1	20	14	US-10-226-007-861	Sequence 861, App
45	35	27.1	20	14	US-10-226-007-864	Sequence 864, App

ALIGNMENTS

RESULT 1
US-09-171-432A-47
; Sequence 47, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Khudiyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171.432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal

FEATURE:

NAME/KEY: Peptide
LOCATION: 1..25
OTHER INFORMATION: /label=YK-1665
US-09-171-432A-47

Query Match 100.0%; Score 129; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRKMGFLF 25
DB 1 QRLKYAQEELSNEVLPPPRKMGFLF 25

RESULT 2

US-09-171-432A-41
Sequence 41, Application US/09171432A
Publication No. US20030187184A1

GENERAL INFORMATION:
APPLICANT: Fields, Howard A.
APPLICANT: Khudyakov, Yuri E.
TITLE OF INVENTION: Antigenically Reactive Regions of the
TITLE OF INVENTION: Hepatitis A Virus Polyprotein
NUMBER OF SEQUENCES: 88

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kilpatrick Stockton LLP
STREET: 3424 Peachtree Road, N.E.
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30326

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

FEATURE:

NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label=YK-1317
US-09-171-432A-41

Query Match 79.8%; Score 103; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 1 QRLKYAQEELSNEVLPPPRK 20

RESULT 3

US-09-171-432A-42
Sequence 42, Application US/09171432A
Publication No. US20030187184A1

GENERAL INFORMATION:
APPLICANT: Fields, Howard A.
APPLICANT: Khudyakov, Yuri E.
TITLE OF INVENTION: Antigenically Reactive Regions of the
TITLE OF INVENTION: Hepatitis A Virus Polyprotein
NUMBER OF SEQUENCES: 88

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kilpatrick Stockton LLP
STREET: 3424 Peachtree Road, N.E.
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30326

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,432A
FILING DATE: 23-NOV-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

FEATURE:

NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label=YK-1318
US-09-171-432A-42

Query Match 50.4%; Score 65; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 VLPPPRKMGFLF 25
DB 1 VLPPPRKMGFLF 12

RESULT 4

US-09-171-432A-40
Sequence 40, Application US/09171432A
Publication No. US20030187184A1

GENERAL INFORMATION:

APPLICANT: Fields, Howard A.

APPLICANT: Khudyakov, Yuri E.

TITLE OF INVENTION: Antigenically Reactive Regions of the

TITLE OF INVENTION: Hepatitis A Virus Polyprotein

NUMBER OF SEQUENCES: 88


```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1316
;
; US-09-171-432A-40
;
; Query Match 34.9%; Score 45; DB 10; Length 20;
; Best Local Similarity 100.0%; Pred. No. 15;
; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 QRLKYAQEE 9
; Db 12 QRLKYAQEE 20
;
; RESULT 5
; US-10-296-734-73
; Sequence 73, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 73
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: The 'Xaa' at location 7 stands for Gln, or Pro.
; FEATURE:
; NAME/KEY: misc_feature
```

```
;
; LOCATION: (11)..(11)
; OTHER INFORMATION: The 'Xaa' at location 11 stands for Gln, or Pro.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: The 'Xaa' at location 16 stands for His, or Leu.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: The 'Xaa' at location 20 stands for Ser, or Leu.
; FEATURE:
; OTHER INFORMATION: GAG segment 32
; US-10-296-734-73
```

```
Query Match 29.5%; Score 38; DB 12; Length 30;
Best Local Similarity 42.1%; Pred. No. 2.4e+02;
Matches 8; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
```

```
Qy 7 QEELSNEVLPPPRKMKGLF 25
Db 9 QEXKDKEXYPPXASLKSLF 27
```

```
RESULT 6
; US-10-097-175-3
; Sequence 3, Application US/10097175
; Publication No. US20030045680A1
; GENERAL INFORMATION:
; APPLICANT: JOYAL, JOHN L.
; APPLICANT: MUELLER, JOHN
; APPLICANT: OZA, VIBHA B.
; APPLICANT: FINDELS, MARK A.
; TITLE OF INVENTION: PEPTIDIC MODULATORS OF THE ANDROGEN RECEPTOR
; FILE REFERENCE: PPI-110
; CURRENT APPLICATION NUMBER: US/10/097,175
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,240
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/352,399
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Androgen Receptor Binding Polypeptides
; US-10-097-175-3
```

```
Query Match 28.7%; Score 37; DB 14; Length 21;
Best Local Similarity 44.4%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy 8 EELSNEVLPPPRKMKGLF 25
Db 4 ESSSDEEDPPAKRKAIF 21
```

```
RESULT 7
; US-10-296-734-936
; Sequence 936, Application US/10296734
; Publication No. US20040054137A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Scott A
; APPLICANT: Ramshaw, Ian A
; TITLE OF INVENTION: Synthetic molecules and uses therefor
; FILE REFERENCE: Savine
; CURRENT APPLICATION NUMBER: US/10/296,734
; CURRENT FILING DATE: 2003-08-04
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 1507
```

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 936
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: MART segment 7
US-10-296-734-936

Query Match      28.7%  Score 37; DB 12; Length 30;
Best Local Similarity 61.3%; Pred. No. 3.3e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      6 AQEELSNEVLPPP 18
        |||||
Db      15 AYEKLSAEQSPPP 27

RESULT 8
US-09-933-767-856
; Sequence 856, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: PCT/US01/05614
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,836
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/193,170
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/205,258
; PRIOR FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: PCT/US98/11422
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/048,885
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,375
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,881
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,880
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,896
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,020
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,876
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,895
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,884
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,894
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,971
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,882
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,899
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,893
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,900
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,901
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,892
```

```
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,915
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,019
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,972
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,916
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,373
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,875
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,374
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,917
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,949
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,883
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,897
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,898
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,962
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,963
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,877
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,878
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 856
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-933-767-856
```

```
Query Match      27.9%; Score 36; DB 10; Length 27;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 15 LPPPRKMKGL 24
DB 16 VPPPLKMPGV 25
      :||| ||| :
      :||| ||| :

RESULT 9
US-10-004-860-856
; Sequence 856, Application US/10004860
; Publication No. US20030065160A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/10/004,860
; CURRENT FILING DATE: 2001-12-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 856
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-004-860-856

Query Match      27.9%; Score 36; DB 12; Length 27;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 15 LPPPRKMKGL 24
DB 16 VPPPLKMPGV 25
      :||| ||| :
      :||| ||| :

RESULT 10
US-10-023-282-856
; Sequence 856, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 856
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-282-856

Query Match      27.9%; Score 36; DB 14; Length 27;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 15 LPPPRKMKGL 24
DB 16 VPPPLKMPGV 25
      :||| ||| :
      :||| ||| :
```



```

; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 840
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Human herpesvirus 4
US-10-226-007-834

```

```

Query Match      27.1%; Score 35; DB 14; Length 14;
Best Local Similarity 58.3%; Pred.No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      7 OEELSNELVLP 18
Db      2 EERESNEEPP 13

```

```

RESULT 15
US-10-226-007-834
; Sequence 834, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 834
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Human herpesvirus 4
US-10-226-007-834

```

```

Query Match      27.1%; Score 35; DB 14; Length 15;
Best Local Similarity 58.3%; Pred.No. 3e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      7 OEELSNELVLP 18
Db      1 EERESNEEPP 12

```

Search completed: July 15, 2004, 15:22:58
Job time : 40.6667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:40:36 ; Search time 10.6667 Seconds
(without alignments)
225.449 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLXVQEELSNEVLPPPRKMGGLP 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 6282

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	25.6	24	2	PC2001
2	30	23.3	30	2	H95021
3	29	22.5	15	2	PT0037
4	29	22.5	27	2	S28940
5	28	21.7	14	2	C33098
6	28	21.7	18	2	S54272
7	28	21.7	22	1	MXKN2
8	28	21.7	24	2	D53288
9	27.5	21.3	30	2	A49955
10	27	20.9	12	2	B39690
11	27	20.9	18	1	DRUFPD
12	27	20.9	18	2	A45590
13	27	20.9	20	2	A42267
14	27	20.9	30	2	H94074
15	26	20.2	18	2	A54195
16	26	20.2	19	2	G56819
17	26	20.2	19	2	S02808
18	26	20.2	21	2	A53733
19	26	20.2	24	2	I39289
20	26	20.2	24	2	G85602
21	26	20.2	26	2	S10614
22	26	20.2	26	2	A28108
23	26	20.2	26	4	A56840
24	26	20.2	27	2	C44524
25	26	20.2	27	2	S35595
26	26	20.2	27	2	JS0676
27	26	20.2	28	2	JQ1035
28	26	20.2	30	2	A05315
29	25	19.4	11	2	D45900

30	25	19.4	12	2	PN0663
31	25	19.4	14	2	S65392
32	25	19.4	15	2	S67918
33	25	19.4	20	2	A45806
34	25	19.4	20	2	JP0070
35	25	19.4	22	1	MXKN1
36	25	19.4	22	1	MXKN3
37	25	19.4	22	2	S58433
38	25	19.4	22	2	JP0071
39	25	19.4	24	2	S70329
40	25	19.4	25	2	S07250
41	25	19.4	27	2	I49747
42	25	19.4	27	2	JT0517
43	25	19.4	28	2	TI4905
44	24	18.6	29	2	T37120
45	24	18.6	15	2	A26212

ALIGNMENTS

RESULT 1

PC2001
major allergen - European chestnut (fragment)
N:Alternate names: Cas sl
C:Species: Castanea sativa (European chestnut)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Mar-1999
C:Accession: PC2001
R:Kos, T.; Hoffmann-Sommergruber, K.; Ferreira, F.; Hirschwehr, R.; Ahorn, H.; Horak, F.;
Biochem. Biophys. Res. Commun. 196, 1086-1092, 1993
A:Title: Purification, characterization and N-terminal amino acid sequence of a new major
A:Reference number: PC2001; MUID:94071857; PMID:7504464
A:Accession: PC2001
A:Molecule type: protein
A:Residues: 1-24 <KOS>
C:Comment: This protein is antigenically related to the major birth pollen allergen Bet v
C:Superfamily: pathogenesis-related protein
C:Keywords: pollen

Query Match 25.6%; Score 33; DB 2; Length 24;
Best Local Similarity 33.3%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 5 YAQEELSNEVLPPPRKMK 22
DB 3 FTSESQTSVIPPARLFX 20

RESULT 2

H95021
hypothetical protein SP0188 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95021
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide
on, J.D.; Maynam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, I.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95021
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-30 <KUR>
A:Cross-references: GB:A5005672; PIDN:AAK74369.1; PID:G14971656; GSPDB:GN00164; TIGR:SP45
A:Experimental source: Strain TIGR4
C:Genetics:
A:Gene: SP0188

Query Match 23.3%; Score 30; DB 2; Length 30;
Best Local Similarity 42.1%; Pred. No. 6.1e+02;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 RLKYAQEELSNEVLPPPK 20
| | | : | : | |
Db 3 RKIYENDEKSQKLKIGRK 21

RESULT 3
PT0037
light harvesting complex chain IIII/b, photosystem I - rice (fragment)
C:Species: Oryza sativa (rice)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1993
C:Accession: PT0037; PS0205
R:Uchiyama, Y.; Tsugita, A.
submitted to JIPID, June 1991
A:Reference number: PS0189
A:Accession: PT0037
A:Molecule type: protein
A:Residues: 1-15 <UCH>

Query Match 22.5%; Score 29; DB 2; Length 15;
Best Local Similarity 46.2%; Pred. No. 3.9e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 AQEELSNEVLPPP 18
| | | : | | |
Db 1 AEEEAAPPPPPP 13

RESULT 4
S28940
cyclic nucleotide-binding phosphatase - potato
C:Species: Solanum tuberosum (potato)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 17-Mar-1999
C:Accession: S28940
R:Polya, G.M.; Wettenhall, R.E.H.
Biochim. Biophys. Acta 1159, 179-184, 1992
A:Title: Rapid purification and N-terminal sequencing of a potato tuber cyclic nucleotid
A:Reference number: S28940; MUID:93003376; PMID:11382614
A:Accession: S28940
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-27 <POL>

Query Match 22.5%; Score 29; DB 2; Length 27;
Best Local Similarity 41.2%; Pred. No. 7.7e+02;
Matches 7; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 RLKYAQEELSNEVLPPP 18
| | | : | | |
Db 10 RLAVETNNLQNSKLVP 26

RESULT 5
C33098
22k exoantigen - malaria parasite (Plasmodium falciparum) (fragments)
C:Species: Plasmodium falciparum
C:Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000
C:Accession: C33098
R:Nichols, J.H.; Hager, L.P.
submitted to the Protein Sequence Database, May 1990
A:Reference number: A33098
A:Accession: C33098
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-14 <NIC>

Query Match 21.7%; Score 28; DB 2; Length 14;
Best Local Similarity 62.5%; Pred. No. 5.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 14 VLPPPRKM 21
| | | : | | |
Db 1 VLFPQEK 8

RESULT 6
S54272

CTC 75 protein - human
C:Species: Homo sapiens (man)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S54272
R:Genersch, E.; Eckerskorn, C.; Lottspeich, F.; Herzog, C.; Kuehn, K.; Poeschl, E.
EMBO J. 14, 791-800, 1995
A:Title: Purification of the sequence-specific transcription factor CTCBF, involved in t
A:Reference number: S54272; MUID:95188883; PMID:7882962
A:Accession: S54272
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-18 <GEN>

Query Match 21.7%; Score 28; DB 2; Length 18;
Best Local Similarity 44.4%; Pred. No. 6.9e+02;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLKYAQEEL 10
| | | : | | |
Db 9 KVEYSBEEL 17

RESULT 7
MXKN2

mu-conotoxin GIIIB [validated] - cone shell (Conus geographus)
N:Alternate names: geographutoxin II (GTx II); myotoxin II
C:Species: Conus geographus (geography cone)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 15-Sep-2000
C:Accession: A01787; B23579
R:Sato, S.; Nakamura, H.; Ohizumi, Y.; Kobayashi, J.; Hirata, Y.
FEBS Lett. 155, 277-280, 1983
A:Title: The amino acid sequences of homologous hydroxyproline-containing myotoxins from
A:Reference number: A91309; MUID:83210170; PMID:6852238
A:Accession: A01787
A:Molecule type: protein
A:Residues: 1-22 <SAT>
R:Cruz, L.J.; Gray, W.R.; Olivera, B.M.; Zeikus, R.D.; Kerr, L.; Yoshikami, D.; Moczydl
J. Biol. Chem. 260, 9280-9288, 1985
A:Title: Conus geographus toxins that discriminate between neuronal and muscle sodium ch
A:Reference number: A23579; MUID:85261316; PMID:2410412
A:Accession: B23579
A:Molecule type: protein
A:Residues: 1-22 <CRU>

R:Hall, J.M.; Alewood, P.F.; Craik, D.J.
submitted to the Brookhaven Protein Data Bank, April 1996
A:Reference number: A65705; PDB:1GIB
A:Contents: annotation; conformation by (1)H-NMR, residues 1-22
R:Hall, J.M.; Alewood, P.F.; Craik, D.J.
Biochemistry 35, 8824-8835, 1996
A:Title: Three-dimensional solution structure of mu-conotoxin GIIIB, a specific blocker
A:Reference number: A58590; MUID:96280640; PMID:8698418
A:Contents: annotation; conformation by (1)H-NMR
C:Superfamily: mu-conotoxin
C:Keywords: amidated carboxyl end; hydroxyproline; myotoxin; sodium channel inhibitor; v

Query Match 21.7%; Score 28; DB 1; Length 22;
Best Local Similarity 83.3%; Pred. No. 8.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 PPRKMK 22
| | | : | | |
Db 6 PPRKCK 11

RESULT 8

D53288
major pollen allergen Que a I - white oak (fragment)
C:Species: Quercus alba (white oak)
C>Date: 02-May-1994 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: D53288
R:Ipse, H.; Hansen, O.C.
Mol Immunol. 28, 1279-1288, 1991
A>Title: The NH2-terminal amino acid sequence of the immunochemically partial identical
s) Car b I and oak (Quercus alba) Que a I pollens.
A:Reference number: A53288; MUID:92072607; PMID:1961201
A:Accession: D53288
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <IPS>
A:Cross-references: PID:G239736; PIDN:AAB20454.1
A:Experimental source: pollen
A>Note: sequence extracted from NCBI backbone (NCBIP:68410)
C:Superfamily: pathogenesis-related protein
C:Keywords: pollen

Query Match 21.7%; Score 28; DB 2; Length 24;
Best Local Similarity 28.6%; Pred. No. 9.5e+02;
Matches 6; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 5 YAQEELSNEVLPPPRKMGKLF 25
DB 3 FTXESQETSVIAPAXLFXALF 23

RESULT 9
A49955
protein-tyrosine kinase (EC 2.7.1.112) ZAP-70 (alternatively spliced form) - human (frag
C:Species: Homo sapiens (man)
C>Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 27-Oct-2003
C:Accession: A49955
R:Arpaia, E.; Shahr, M.; Dadi, H.; Cohen, A.; Roifman, C.M.
Cell 76, 947-958, 1994
A>Title: Defective T cell receptor signaling and CD8+ thymic selection in humans lacking
A:Reference number: A49955; MUID:94170394; PMID:8124727
A:Accession: A49955
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-30 <ARP>
A:Experimental source: selective T-cell deficiency, patient P1F1, peripheral blood monon
A>Note: sequence extracted from NCBI backbone (NCBIP:145660)
C:Genetics:
A:Gene: GDB:SRK; ZAP-70
A:Cross-references: GDB:433738; OMIM:176947
A:Map position: 4pter-4qter
A>Note: defects in this gene are associated with an autosomal recessive form of severe c
C:Superfamily: protein-tyrosine kinase SYK; protein kinase homology; SH2 homology
C:Keywords: ATP; phosphotransferase; immunodeficiency; T-cell; tyrosine-specific protein

Query Match 21.3%; Score 27.5; DB 2; Length 30;
Best Local Similarity 34.8%; Pred. No. 1.5e+03;
Matches 8; Conservative 3; Mismatches 7; Indels 5; Gaps 1;

QY 1 QRLKYAQEELSNEVLPPPRKMGK 23
DB 11 BALSYGQKPYKLE----QKMGK 28

RESULT 10
B39690
neural cell adhesion molecule, cardiac splice form +,-,- - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 16-Jul-1999
C:Accession: B39690
R:Reyes, A.A.; Small, S.J.; Akesson, R.
Mol. Cell. Biol. 11, 1654-1661, 1991
A>Title: At least 27 alternatively spliced forms of the neural cell adhesion molecule mR
A:Reference number: A39690; MUID:91141516; PMID:1996115
A:Accession: B39690

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-12 <REY>
A:Cross-references: GB:M63970
C:Keywords: cardiac muscle; cell adhesion; heart

Query Match 20.9%; Score 27; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 6.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 PPPRX 20
DB 6 PPPRE 10

RESULT 11
DRUFFD
pigment-dispersing hormone - Atlantic sand fiddler crab
N:Alternate names: PDH
C:Species: Uca pugnator (Atlantic sand fiddler crab)
C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 08-Dec-1995
C:Accession: A25144
R:Rao, K.R.; Riehm, J.P.; Zahn, C.A.; Kleinholtz, L.H.; Tarr, G.E.; Johnson, L.; Norton,
Proc. Natl. Acad. Sci. U.S.A. 82, 5319-5322, 1985
A>Title: Characterization of a pigment-dispersing hormone in eyestalks of the fiddler cra
A:Reference number: A25144
A:Accession: A25144
A:Molecule type: protein
A:Residues: 1-18 <RAO>
C:Superfamily: pigment-dispersing hormone
C:Keywords: amidated carboxyl end; neuropeptide
F18/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 20.9%; Score 27; DB 1; Length 18;
Best Local Similarity 46.2%; Pred. No. 9.7e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 9 ELSNEVLPPPRKM 21
DB 3 ELINSILGLPKVM 15

RESULT 12
A45590
beta-pigment-dispersing hormone analog - red swamp crayfish
C:Species: Procambarus clarkii (red swamp crayfish)
C>Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 11-Apr-1995
C:Accession: A45590
R:McCallum, M.L.; Rao, K.R.; Riehm, J.P.; Mohr, C.J.; Morgan, W.T.
Pigment Cell Res. 4, 201-208, 1991
A>Title: Primary structure and relative potency of an analog of beta-PDH (pigment-dispers
A:Reference number: A45590; MUID:92390305; PMID:1823925
A:Accession: A45590
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-18 <MCC>
A:Experimental source: eyestalk
A>Note: sequence extracted from NCBI backbone (NCBIP:112439)

Query Match 20.9%; Score 27; DB 2; Length 18;
Best Local Similarity 46.2%; Pred. No. 9.7e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 9 ELSNEVLPPPRKM 21
DB 3 ELINSILGLPKVM 15

RESULT 13
A42267
J-kappa recombination sequence-binding protein RBP-2N - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999

Search completed: July 15, 2004, 14:52:15
Job time : 11.6667 secs

C;Accession: A42267; S23798
R;Kawauchi, M.; Oka, C.; Shibayama, S.; Koromilas, A.E.; Matsunami, N.; Hamaguchi, Y.; H
J. Biol. Chem. 267, 4016-4022, 1992
A;Title: Genomic organization of mouse J-kappa recombination signal binding protein (RBE
A;Reference number: A42267; MUID:92156146; PMID:1740450
A;Accession: A42267
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-20 <KAW>
A;Cross-references: GB:M01866

Query Match 20.9%; Score 27; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 16 PPPRRM 21
|||:
Db 13 PPPKRL 18

RESULT 14

H84074
Hypothetical protein BH3400 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: H84074
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84074
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-30 <STO>
A;Cross-references: GB:AP001518; GB:BA000004; NID:gi0175792; PIDN:BA07119.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3400

Query Match 20.9%; Score 27; DB 2; Length 30;
Best Local Similarity 31.2%; Pred. No. 1.7e+03;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QRLKYAOBELSNEVLP 16
|||:
Db 2 QSLILVKQTLNKMP 17

RESULT 15

A54195
Na+/K+-exchanging ATPase (EC 3.6.3.9) - spiny dogfish (fragment)
C;Species: Squalus acanthias (spiny dogfish)
C;Date: 13-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 19-Apr-2002
C;Accession: A54195
R;Bmann, M.; Karlisch, S.J.; Sottrup-Jensen, L.; Marsh, D.
Biochemistry 33, 8044-8050, 1994
A;Title: Structural integrity of the membrane domains in extensively trypsinized Na,K-AT
A;Reference number: A54195; MUID:94297020; PMID:8025109
A;Accession: A54195
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <ESN>
A;Experimental source: rectal gland
A;Note: sequence extracted from NCBI backbone (NCBIP:149363)
C;Keywords: hydrolase

Query Match 20.2%; Score 26; DB 2; Length 18;
Best Local Similarity 57.1%; Pred. No. 1.4e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 NEVLPPP 18
|||:
Db 4 NALTPPP 10

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:01 ; Search time 6.6667 Seconds
(without alignments)
195.263 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLKYAQEELSNEVLPFRKMGKLF 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 2073

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	24.0	13	1	PEDI_HYDAT
2	28	21.7	17	1	APID_BOMPA
3	28	21.7	22	1	CXM2_CONGE
4	28	21.7	25	1	ACP_RICCA
5	27	20.9	13	1	FIBB_RABIT
6	27	20.9	18	1	PDH_PROCL
7	27	20.9	18	1	PDH_UCAPU
8	26	20.2	19	1	PSAE_CUCSA
9	26	20.2	24	1	LPER_STRFR
10	26	20.2	26	1	RL16_BACST
11	26	20.2	27	1	DCUP_RHOSH
12	25	19.4	18	1	PDHC_ORCLT
13	25	19.4	22	1	CXM1_CONGE
14	25	19.4	22	1	CXM3_CONGE
15	24.5	19.0	29	1	RP54_CLOKL
16	24	18.6	15	1	CPBP_PROAT
17	24	18.6	16	1	IBP4_PIG
18	24	18.6	20	1	PSAF_MAIZE
19	24	18.6	21	1	SCIB_BPTS
20	24	18.6	23	1	TL17_SPTOL
21	24	18.6	28	1	SMS2_ORENI
22	23.5	18.2	30	1	VTTA_BPT3
23	23	17.8	15	1	PRP_MYCBO
24	23	17.8	16	1	FOR1_MYRGU
25	23	17.8	18	1	PDHB_ORCLT
26	23	17.8	18	1	PDH_PANBO
27	23	17.8	20	1	THIO_CANFA
28	23	17.8	21	1	SYA_RAT
29	23	17.8	23	1	CR41_LITCE
30	23	17.8	23	1	CR42_LITCE
31	23	17.8	23	1	CR43_LITCE
32	23	17.8	25	1	ATPD_MICLU
33	23	17.8	26	1	HR83_TRITO

34	23	17.8	29	1	GLUC_CHLBR	P31297 chinchilla
35	23	17.8	30	1	ACB1_DIGLA	P81624 digitalis l
36	22	17.1	11	1	TKNA_HORSE	P01290 equus cabal
37	22	17.1	11	1	TKNA_RANRI	P29207 rana ridibu
38	22	17.1	12	1	PKK4_PERFU	P82690 periplaneta
39	22	17.1	20	1	FIBB_ELEMA	P14538 elephas max
40	22	17.1	22	1	BS23_SERPL	P83378 serratia pl
41	22	17.1	22	1	DEF1_SPTOL	P81572 serratia ol
42	22	17.1	22	1	HELN_HELVI	P83427 heliothis v
43	22	17.1	22	1	PSBQ_ORYSA	P83646 oryza sativ
44	22	17.1	23	1	NIFD_ANASL	P31177 anabaena sp
45	22	17.1	23	1	UDP_LACRH	P19662 lactobacill

ALIGNMENTS

RESULT 1
PEDI_HYDAT
ID_PEDI_HYDAT STANDARD; PRT; 13 AA.
AC P80578;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Pedin.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE.
RX MEDLINE=96232307; PubMed=8674432;
RA Hoffmeister S.A.H.;
RT "Isolation and characterization of two new morphogenetically active peptides from Hydra vulgaris.";
RL Development 122:1941-1948(1996).
CC -!- FUNCTION: Morphogenetically active peptide. Active in foot development.
KW Morphogen.
SQ SEQUENCE 13 AA; 1512 MW; 6F5266E3F37632CB CRC64;
Query Match 24.0%; Score 31; DB 1; Length 13;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 EELSNEVLP 16
Db 1 EELRPEVLP 9
||| |||
||| |||

RESULT 2
APID_BOMPA
ID APID_BOMPA STANDARD; PRT; 17 AA.
AC P81464;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Apidaecin.
OS Bombus pascuorum (Brown bumble bee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Apoidea;
OC Apidae; Bombus.
OX NCBI_TaxID=65598;
RN [1]
RP SEQUENCE.
RX TISSUE=Hemolymph;
RC MEDLINE=97362903; PubMed=9219367;
RA Rees J.A.; Moniatte M.; Bulet P.;
RT "Novel antibacterial peptides isolated from a European bumblebee, Bombus pascuorum (Hymenoptera, Apoidea).";
RL Insect Biochem. Mol. Biol. 27:413-422(1997).
CC -!- FUNCTION: Antibacterial peptide active against Gram-negative bacteria.

CC -1- INDUCTION: BY bacterial infection.
 DR InterPro: IPR004828; Apidaecin.
 DR Pfam: PF00807; Apidaecin; 1.
 KW Insect immunity; Antibiotic; Hemolymph.
 SQ SEQUENCE 17 AA; 1963 MW; CDD0D02C8BC23D1 CRC64;

Query Match 21.7%; Score 28; DB 1; Length 17;
 Best Local Similarity 80.0%; Pred. No. 4.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 LPPPR 19
 :|||
 Db 7 IPPPR 11

RESULT 3

CX2 CONGE STANDARD; PRT; 22 AA.
 AC P01524;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mu-conotoxin GIIIB (Myotoxin II) (Geographutoxin II) (GTX-II).
 OS Conus geographus (Geography cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6491;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85261316; PubMed=2410412;
 RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,
 RA Yoshikami D., Moczydlowski E.;
 RT "Conus geographus toxins that discriminate between neuronal and
 muscle sodium channels.";
 RL J. Biol. Chem. 260:9280-9288 (1985).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=83210170; PubMed=6852238;
 RA Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;
 RT "The amino acid sequences of homologous hydroxyproline-containing
 myotoxins from the marine snail Conus geographus venom.";
 RL FEBS Lett. 155:277-280 (1983).
 RN [3]

RP STRUCTURE BY NMR.
 RX MEDLINE=96280640; PubMed=8688418;
 RA Hill J.M., Alewood P.F., Craik D.J.;
 RT "Three-dimensional solution structure of mu-conotoxin GIIIB, a
 specific blocker of skeletal muscle sodium channels.";
 RL Biochemistry 35:8824-8835 (1996).
 CC -1- FUNCTION: Mu-conotoxins bind and block voltage-sensitive sodium
 channel (VSSC).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -1- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS. MU-TYPE
 FAMILY.

DR PIR; A01787; MXKN2.
 DR PDB; 1GIB; 08-NOV-96.
 DR InterPro: IPR008036; Mu-conotoxin.
 DR Pfam; PF05374; Mu-conotoxin; 1.
 KW Toxin; Sodium channel inhibitor; Hydroxylation; Amidation;
 XW 3D-structure.
 FT DISULFID 3 15
 FT DISULFID 4 20
 FT DISULFID 10 21
 FT MOD_RES 6 6 HYDROXYLATION.
 FT MOD_RES 7 7 HYDROXYLATION.
 FT MOD_RES 17 17 HYDROXYLATION.
 FT MOD_RES 22 22 AMIDATION.
 SQ SEQUENCE 22 AA; 2599 MW; F50402BA93199E01 CRC64;

Query Match 21.7%; Score 28; DB 1; Length 22;
 Best Local Similarity 83.3%; Pred. No. 5.4e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 17 PPRMK 22
 :|||
 Db 6 PPRCK 11

RESULT 4

ACP ACICA STANDARD; PRT; 25 AA.
 AC P80916;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acyl carrier protein (ACP) (Fragment).
 GN ACP.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=ATCC 33604;
 RX MEDLINE=97315245; PubMed=9171419;
 RA Tang L., Weissborn A.C., Kennedy E.P.;
 RT "Domains of Escherichia coli acyl carrier protein important for
 membrane-derived-oligosaccharide biosynthesis.";
 RL J. Bacteriol. 179:3697-3705 (1997).
 CC -1- FUNCTION: Carrier of the growing fatty acid chain in fatty
 acid biosynthesis.
 CC -1- PATHWAY: De novo fatty acid biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- PTM: 4'-phosphopantetheine is transferred from CoA to a specific
 serine of apo-ACP by acps. This modification is essential for
 activity because fatty acids are bound in thioester linkage to the
 sulphydryl of the prosthetic group.
 CC -1- SIMILARITY: Contains 1 acyl carrier domain.
 DR HAMAP: MF 01217; -; 1.
 DR InterPro: IPR006163; Pp_bind.
 DR InterPro: IPR006162; Ppantne_S.
 DR PROSITE; PS00012; PHOSPHOPANTHEINE; PARTIAL.
 DR PROSITE; PS50075; ACP DOMAIN; PARTIAL.
 KW Lipid synthesis; Fatty acid biosynthesis; Phosphopantetheine.
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 2886 MW; 38F186AC1C7A3907 CRC64;

Query Match 21.7%; Score 28; DB 1; Length 25;
 Best Local Similarity 38.1%; Pred. No. 6.2e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 8; Gaps 1;

QY 1 ORLYA-----QEELSNE 13
 :|||
 Db 5 QRIQVAEQLGWRAEIKNE 25

RESULT 5

FIBB RABIT STANDARD; PRT; 13 AA.
 AC P14478;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE.
 RA Blomback B., Blomback M., Grondahl N.J.;
 RT "Studies on fibrinopeptides from mammals.";
 RL Acta Chem. Scand. 19:1789-1791 (1965).

```

CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
CC InterPro: IPR002181; Fibrinogen C.
DR PROSITE; PS00514; FIBRIN AG C DOMAIN; PARTIAL.
KW Blood coagulation; Plasma; Sulfation.
FT PRETIDE 1 13 FIBRINOPEPTIDE B.
FT MOD_RES 4 4 SULFATION.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1493 MW; 9417F3B76337CB1A CRC64;

Query Match 20.9%; Score 27; DB 1; Length 13;
Best Local Similarity 41.7%; Pred. No. 4.3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 3; Mismatches 4;

Qy 8 ELSNEVLPPPR 19
Db 2 DDYDDEVLPDAR 13

RESULT 6
PDH PROCL
ID - PDH PROCL STANDARD; PRT; 18 AA.
AC QPTW7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pigment-dispersing hormone (PDH).
OS Procambatus clarkii (Red swamp crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Astacoidea; Cambaridae; Procambatus.
OX NCBI_TaxID=6728;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Eyestalk;
RX MEDLINE=92390305; PubMed=1823925;
RA McCallum M.L., Rao K.R., Riehm J.P., Mohrerr C.J., Morgan W.T.;
RT "Primary structure and relative potency of an analog of beta-PDH
RL (pigment-dispersing hormone) from the crayfish Procambatus clarkii.";
RL Pigment Cell Res. 4:201-208(1991).
CC -!- FUNCTION: Causes the migration of the distal retinal pigment into
CC the proximal end of the pigment chromatophore cells and thus
CC decreases the amount of light entering the retinulas.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC DR PIR; A45590; A45590.
KW Hormone; Amidation.
FT MOD_RES 18 18 AMIDATION.
FT SEQUENCE 18 AA; 1942 MW; 25D47E8D016F544E CRC64;

Query Match 20.9%; Score 27; DB 1; Length 18;
Best Local Similarity 46.2%; Pred. No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 5;

Qy 9 ELSNEVLPPPRKM 21
Db 3 ELINSILGLPKVM 15

RESULT 7
PDH UCAPU
ID - PDH UCAPU STANDARD; PRT; 18 AA.
AC P08871;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

```

```

DE Pigment-dispersing hormone (PDH) (Beta-PDH) (Light adapting distal
DE retinal pigment hormone) (DRPH).
OS Uca pugnator (Atlantic sand fiddler crab) (Celuca pugnator).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Ocypodoidea; Ocypodidae; Ocypodinae; Uca complex;
OC Celuca.
OX NCBI_TaxID=6772;
RN [1]
RP SEQUENCE.
RC TISSUE=Eyestalk;
RA Rao K.R., Riehm J.P., Zahnaw C.A., Kleinholz L.H., Tarr G.E.,
RA Johnson L., Norton S., Landau M., Semmes O.J., Sattelberg R.M.,
RA Jorenbey W.H., Hintz M.F.;
RA "Characterization of a pigment-dispersing hormone in eyestalks of the
RT fiddler crab Uca pugnator.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:5319-5322(1985).
RN [2]
RP SEQUENCE.
RX MEDLINE=93230895; PubMed=8472537;
RA Loehr J., Klein J., Webster S.G., Dirksen H.;
RT "Quantification, immunoaffinity purification and sequence analysis of
RT a pigment-dispersing hormone of the shore crab, Carcinus maenas
RT (L.).";
RL Comp. Biochem. Physiol. 104B:699-706(1993).
CC -!- FUNCTION: Causes the migration of the distal retinal pigment into
CC the proximal end of the pigment chromatophore cells and thus
CC decreases the amount of light entering the retinulas.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC DR PIR; A25144; DRUFPD.
KW Hormone; Amidation.
FT MOD_RES 6 18 AMIDATION.
FT SEQUENCE 18 AA; 1928 MW; 25D5CE8D016F544E CRC64;

Query Match 20.9%; Score 27; DB 1; Length 18;
Best Local Similarity 46.2%; Pred. No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 5;

Qy 9 ELSNEVLPPPRKM 21
Db 3 ELINSILGLPKVM 15

RESULT 8
PSAE CUCSA
ID - PSAE CUCSA STANDARD; PRT; 19 AA.
AC P42047;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Photosystem I reaction center subunit IV (PSI-E) (Photosystem I 19.5
DE kDa protein) (PS I subunit 6) (Fragment).
GN PSAE.
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eumids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX NCBI_TaxID=3659;
RN [1]
RP SEQUENCE.
RC TISSUE=Cotyledon;
RX MEDLINE=91355209; PubMed=1883835;
RA Iwasaki Y., Ishikawa H., Hibino T., Takabe T.;
RT "Characterization of genes that encode subunits of cucumber PS I
RT complex by N-terminal sequencing.";
RL Biochim. Biophys. Acta 1059:141-148(1991).
CC -!- FUNCTION: Stabilizes the interaction between psaC and the PSI
CC core, assists the docking of the ferredoxin to PSI and interacts
CC with ferredoxin-NADP oxidoreductase (By similarity).
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- SIMILARITY: Belongs to the psaE family.

```

```

DR PIR; G56819; G56819.
KW Photosynthesis; Photosystem I; Chloroplast; Thylakoid; Membrane.
FT NON_TER 19
SQ SEQUENCE 19 AA; 1810 MW; A93E8BDD089F738 CRC64;

Query Match 20.2%; Score 26; DB 1; Length 19;
Best Local Similarity 57.1%; Pred. No. 9.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 NEVLPPP 18
Db :|||
2 DEAAAPP 8

RESULT 9
LPER_STRFR STANDARD; PRT; 24 AA.
AC P45440;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Erythromycin resistance leader peptide (23S rRNA methylase leader peptide).
OS Streptomyces fradiae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OX Streptomycineae; Streptomycetaceae; Streptomyces.
ON NCBI_TaxID=1906;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88169508; PubMed=3127381;
RA Kamimura S., Weisblum B.;
RT "Translational attenuation control of ermSF, an inducible resistance determinant encoding rRNA N-methyltransferase from Streptomyces fradiae.";
RL J. Bacteriol. 170:1800-1811(1988).
CC -!- FUNCTION: THIS PEPTIDE IS INVOLVED IN THE CONTROL MECHANISM OF THE SYNTHESIS OF THE MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B RESISTANCE PROTEIN.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-3 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54994; CA38741.1; -.
DR PIR; S10614; S10614.
DR InterPro; IPR000114; Ribosomal L16.
DR PROSITE; PS00586; RIBOSOMAL_L16_1; PARTIAL.
DR PROSITE; PS00701; RIBOSOMAL_L16_2; PARTIAL.
KW Ribosomal protein; rRNA-binding.
FT NON_TER 26
SQ SEQUENCE 26 AA; 3225 MW; 3AF5C568C77B185B CRC64;

Query Match 20.2%; Score 26; DB 1; Length 26;
Best Local Similarity 30.4%; Pred. No. 1.3e+03;
Matches 7; Conservative 4; Mismatches 4; Indels 8; Gaps 1;

QY 1 QRLKYAQEELSNEVLPPPRMKG 23
Db :||| :|||
5 KRVKYRREHRG-----RMKG 19

RESULT 11
DCUP_RHOSH STANDARD; PRT; 27 AA.
AC P32920;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD) (Fragment).
GN HEME.
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OX Rhodobacteraceae; Rhodobacter.
ON NCBI_TaxID=1063;
RN [1]
RP SEQUENCE.
RC STRAIN=ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RX MEDLINE=93356732; PubMed=8352737;
RA Jones R.M., Jordan P.M.;
RT "Purification and properties of the uroporphyrinogen decarboxylase from Rhodobacter sphaeroides.";
RL Biochem. J. 293:703-712(1993).
CC -!- CATALYTIC ACTIVITY: Uroporphyrinogen-III = coproporphyrinogen + 4 CO(2).
CC -!- PATHWAY: Porphyrin biosynthesis.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the uroporphyrinogen decarboxylase family.
DR PIR; S35595; S35595.
DR HAMAP; MF_00218; -.
DR InterPro; IPR000257; Uro decarboxyls.
DR PROSITE; PS00906; UROD_1; PARTIAL.
DR PROSITE; PS00907; UROD_2; PARTIAL.
KW Lyase; decarboxylase; Porphyrin biosynthesis;
KW Chlorophyll biosynthesis.
FT NON_TER 27

```

SQ SEQUENCE 27 AA; 3042 MW; BAF93BD701A29BAA CRC64;
 Query Match 20.2%; Score 26; DB 1; Length 27;
 Best Local Similarity 55.6%; Pred. No. 1.4e+03;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 10 LSNEVLPPPP 18
 | | | | |
 Db 6 LKGETLPTP 14

RESULT 12
 PDHC_ORCLI STANDARD; PRT; 18 AA.
 AC P83587;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pigment-dispersing hormone C (PDH C) (Light adapting distal retinal
 DE pigment hormone C) (DRPH C).
 OS Orconetes limosus (Spinycreek crayfish).
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacoidea;
 CC Astacoidea; Cambaridae; Orconectes.
 OX NCBI_TaxID=28379;
 RN [1]
 RN SEQUENCE, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, MASS SPECTROMETRY,
 RP AND AMIDATION.
 RC TISSUE=Eyestalk;
 RA Bullau P., Schmitz T., Peter-Katalinic J., Keller R.;
 RL Submitted (MAY-2003) to Swiss-Prot.
 CC -!- FUNCTION: Causes the migration of the distal retinal pigment into
 CC the proximal end of the pigment chromatophore cells and thus
 CC decreases the amount of light entering the retinulas. May also
 CC function as a neurotransmitter and/or neuromodulator (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Eyestalk sinus gland.
 CC -!- MASS SPECTROMETRY: MW=1855.92; METHOD=Electrospray.
 CC -!- SIMILARITY: Belongs to the arthropod PDH family.
 KW Hormone; Neuropeptide; Neurotransmitter; Amidation.
 FT MOD_RES 18 18
 FT AMIDATION.
 SQ SEQUENCE 18 AA; 1958 MW; F9C472ABCE9DC45C CRC64;
 Query Match 19.4%; Score 25; DB 1; Length 18;
 Best Local Similarity 46.2%; Pred. No. 1.2e+03;
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 9 ELSNEVLPPPKM 21
 | | | | |
 Db 3 ELINAILGPTLM 15

RESULT 13
 CXM1_CONGE STANDARD; PRT; 22 AA.
 AC P01523;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mu-conotoxin GI1IA (Myotoxin I) (Geographutoxin I) (GTx-I).
 OS Conus geographus (Geography cone).
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 CC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
 CC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6491;
 RN [1]
 RN SEQUENCE.
 RX MEDLINE=85261316; PubMed=2410412;
 RA Cruz L.J., Gray W.R., Olivera B.M., Zeikus R.D., Kerr L.,
 RA Yoshikami D., Moczydlowski E.;
 RT "Conus geographus toxins that discriminate between neuronal and
 DE muscle sodium channels.";

RL J. Biol. Chem. 260:9280-9288 (1985).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=83210170; PubMed=6852238;
 RA Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.;
 RT "The amino acid sequences of homologous hydroxyproline-containing
 RT myotoxins from the marine snail Conus geographus venom.";
 RL FEBS Lett. 155:277-280 (1983).
 RN [3]
 RP DISULFIDE BONDS.
 RX MEDLINE=90249506; PubMed=2338142;
 RA Hidaka Y., Sato K., Nakamura H., Kobayashi J., Ohizumi Y.,
 RA Simonishi Y.;
 RT "Disulfide pairings in geographutoxin I, a peptide neurotoxin from
 RT Conus geographus.";
 RL FEBS Lett. 264:29-32 (1990).
 RN [4]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91122275; PubMed=1991506;
 RA Ott K.-H., Becker S., Gordon R.D., Rueterjans H.;
 RT "Solution structure of mu-conotoxin GI1IA analysed by 2D-NMR and
 RT distance geometry calculations.";
 RL FEBS Lett. 278:160-166 (1991).
 RN [5]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91299744; PubMed=2069951;
 RA Lancelin J.-M., Kohda D., Iate S.-I., Yanagawa Y., Abe T., Satake M.,
 RA Inagaki F.;
 RT "Tertiary structure of conotoxin GI1IA in aqueous solution.";
 RL Biochemistry 30:6908-6916 (1991).
 CC -!- FUNCTION: Mu-conotoxins bind and block voltage-sensitive sodium
 CC channel (VSSC).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS. MU-TYPE
 CC FAMILY.
 CC PIR: A01786; MXXN1.
 DR PDB; 1TCG; 31-JAN-94.
 DR PDB; 1TCH; 31-JAN-94.
 DR PDB; 1TCJ; 31-JAN-94.
 DR PDB; 1TCI; 31-JAN-94.
 DR PDB; 1TCJ; 31-JAN-94.
 DR InterPro; IPR008036; Mu-conotoxin.
 DR Pfam; PF05374; Mu-conotoxin; 1.
 KW Toxin; Sodium channel inhibitor; Hydroxylation; Amidation;
 KW 3D-structure.
 FT DISULFID 3 15
 FT DISULFID 4 20
 FT DISULFID 10 21
 FT MOD_RES 6 6
 FT MOD_RES 7 7
 FT MOD_RES 17 17
 FT MOD_RES 22 22
 FT MOD_RES 13 16
 FT HELIX 19 21
 FT TURN 19 21
 SQ SEQUENCE 22 AA; 2568 MW; F6CB02ADB359813C CRC64;
 Query Match 19.4%; Score 25; DB 1; Length 22;
 Best Local Similarity 66.7%; Pred. No. 1.5e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 17 PPRKMK 22
 | | | | |
 Db 6 PPAKCK 11

RESULT 14
 CXM3_CONGE STANDARD; PRT; 22 AA.
 AC P05482;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Mu-conotoxin GI1IC.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:31 ; Search time 32.1667 Seconds

(without alignments)
245.221 Million cell updates/sec

Title: US-09-171-432A-47

Perfect score: 129

Sequence: 1 QRLKVAQEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 18020

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mbc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	26.4	22	1 Q9UWI9	Q9uw19 methanospirillum
2	33	25.6	24	10 Q9S8Q4	Q9s8q4 castanea sa
3	33	25.6	29	11 Q8C837	Q8c837 mus musculus
4	32	24.8	26	2 Q936D3	Q936d3 staphylococ
5	31	24.0	16	6 Q9TQZ7	Q9tqz7 bos taurus
6	30	23.3	22	6 Q9TQZ7	Q9tnl1 monodelphis
7	30	23.3	23	13 Q9PV93	Q9pv93 gallus gall
8	30	23.3	30	16 Q97SX5	Q97sx5 streptococc
9	29	22.5	26	11 Q63337	Q63337 rattus norv
10	29	22.5	27	10 Q9S8V4	Q9s8v4 solanum tub
11	29	22.5	29	11 Q8K362	Q8k362 mus musculu
12	28	21.7	15	10 P82439	P82439 nicotiana t
13	28	21.7	21	2 Q9S1C0	Q9s1c0 porphyromon
14	28	21.7	21	4 Q9UCC3	Q9ucc3 homo sapien
15	28	21.7	26	15 Q8Q7B9	Q8q7b9 simian t-ly
16	28	21.7	27	5 Q26204	Q26204 phagocata w

17	28	21.7	27	12	O37240	hepatitis c
18	28	21.7	27	12	O37239	hepatitis c
19	28	21.7	27	12	O37244	hepatitis c
20	28	21.7	27	12	O37242	hepatitis c
21	28	21.7	27	12	O37222	hepatitis c
22	28	21.7	27	12	O37241	hepatitis c
23	28	21.7	27	12	O37238	hepatitis c
24	28	21.7	27	12	O37237	hepatitis c
25	28	21.7	27	12	O37243	hepatitis c
26	28	21.7	27	12	O37223	hepatitis c
27	28	21.7	29	8	Q9G5Y5	Q9g5y5 agama impal
28	28	21.7	30	10	O8GT04	O8gtd4 cicor ariet
29	27.5	21.3	30	4	Q9UBS6	Q9ubs6 homo sapien
30	27	20.9	14	11	Q99PB8	Q99pb8 mus musculu
31	27	20.9	16	5	Q9TWN7	Q9twn7 toxoplasma
32	27	20.9	18	4	Q9UCT9	Q9uct9 homo sapien
33	27	20.9	19	2	Q9RBV1	Q9rbv1 pseudomonas
34	27	20.9	20	3	Q9UR76	Q9ur76 candida alb
35	27	20.9	20	12	Q85617	Q85617 reovirus (t
36	27	20.9	22	12	Q85664	Q85664 reovirus (t
37	27	20.9	22	13	Q90ZV4	Q90zv4 rallus limi
38	27	20.9	22	13	Q90ZV2	Q90zv2 sarothrua
39	27	20.9	22	13	Q90ZV3	Q90zv3 rallus long
40	27	20.9	22	13	Q91OC6	Q91oc6 gallinula c
41	27	20.9	25	4	Q8NE43	Q8ne43 homo sapien
42	27	20.9	25	11	Q64135	Q64135 mus sp. cal
43	27	20.9	25	11	Q8CF86	Q8cf86 rattus norv
44	27	20.9	25	15	Q9WSP2	Q9wsp2 human immun
45	27	20.9	26	2	Q9F4H7	Q9f4h7 klebsiella

ALIGNMENTS

RESULT 1
Q9UWI9 PRELIMINARY; PRT; 22 AA.
AC Q9UWI9
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE 5 kDa flagellin (Fragment).
OS Methanospirillum hungatei.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanomicrobiales; Methanospirillaceae; Methanospirillum.
OX NCBI_TaxID=2203;
RN [1]
RP SEQUENCE.
RX MEDLINE=95095917; PubMed=8002572;
RA Faguy D.M., Koval S.F., Jarrell K.F.;
RT "Physical characterization of the flagella and flagellins from
RT Methanospirillum hungatei.";
RL J. Bacteriol. 176:7491-7498 (1994).
SQ SEQUENCE 22 AA; 2360 MW; 168505F27B4F79B5 CRC64;

Query Match 26.4%; Score 34; DB 1; Length 22;
Best Local Similarity 46.2%; Pred. No. 3.9e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 LSNEVLPPPRKMK 22
Db 3 LTNBILVTEKVK 15

RESULT 2
Q9S8Q4 PRELIMINARY; PRT; 24 AA.
ID Q9S8Q4
AC Q9S8Q4
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE CAS S l major allergen (Fragment).
OS Castanea sativa (Sweet chestnut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eusoids I; Fagales; Fagaceae; Castanea.
OX NCBI_TaxID=21020;

RP SEQUENCE
[1]

RX MEDLINE=94071857; PubMed=7504464;
RA Kos T., Hoffmann-Sommergruber K., Ferreira F., Hirschwehr R.,
RA Ahorn H., Horak F., Jager S., Sperr W., Kraft D., Scheiner O.;
RT "Purification, characterization and N-terminal amino acid sequence of
RT a new major allergen from European chestnut pollen--Cas s 1.";
RL Biochem. Biophys. Res. Commun. 196:1086-1092(1993).
CC -1- SIMILARITY: BELONGS TO THE BETVI FAMILY OF PATHOGENESIS-RELATED
CC PROTEINS.

DR PIR; PC2001.
DR HSSP; O24248; IE09.

DR GO; GO:0009607; P:response to biotic stimulus; IEA.
KW Pathogenesis-related protein; Plant defense.

SQ SEQUENCE 24 AA; 2674 MW; 8FC9A056705F086 CRC64;

Query Match 25.6%; Score 33; DB 10; Length 24;
Best Local Similarity 33.3%; Pred. No. 6.4e+02;

Matches 6; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

OY 5 YAOEELSNEVLPPRKMK 22
Db 3 FTSESQTSVPPARLFK 20

RESULT 3

ID ORC837 PRELIMINARY; PRT; 29 AA.

AC QRC837;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE SLM-1 (Fragment).

GN KHDRBS2 OR 6330586C16RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).

DR ENBL; AK048524; BAC33360.1; -.
DR MGD; MGI:2159649; KHDRBS2.

FT NON TER 1
SQ SEQUENCE 29 AA; 3348 MW; 280CC988D424925 CRC64;

Query Match 25.6%; Score 33; DB 11; Length 29;
Best Local Similarity 62.5%; Pred. No. 7.4e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 16 PPRKMKMG 23
Db 12 PPRRSARG 19

RESULT 4

ID Q936D3 PRELIMINARY; PRT; 26 AA.

AC Q936D3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Ribonucleotide-diphosphate reductase alpha chain (EC 1.17.4.1)

DE (Fragment).

GN NRDE.

OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=1280;
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RN4220;

RA Masahisa M., Borovok I., Schreiber R., Aharonowitz Y., Cohen G.;
RT "Transcriptional analysis of the Staphylococcus aureus aerobic class
RT Ib an anaerobic class III ribonucleotide reductase genes in response
RT to oxygen.";
RL J. Bacteriol. 183:0-0(2001).

DR ENBL; AJ312387; CAC43040.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.

KW Oxidoreductase.

FT NON TER 26
SQ SEQUENCE 26 AA; 3241 MW; D3C2A97CA637EC3A CRC64;

Query Match 24.8%; Score 32; DB 2; Length 26;
Best Local Similarity 63.6%; Pred. No. 9.4e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 KYAQEELSNEV 14
Db 8 KYNHIELNNEV 18

RESULT 5

ID Q9TQZ7 PRELIMINARY; PRT; 16 AA.

AC Q9TQZ7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE Factor H (Fragment).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]

RP SEQUENCE
RX MEDLINE=96202005; PubMed=8615824;

RA Soames C.J., Day A.J., Sim R.B.;
RT "Prediction from sequence comparisons of residues of factor H involved
RT in the interaction with complement component C3b.";
RL Biochem. J. 315:523-531(1996).

SQ SEQUENCE 16 AA; 1871 MW; 7052751E485EC825 CRC64;

Query Match 24.0%; Score 31; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 PPRK 20
Db 6 PPRK 10

RESULT 6

ID Q9TTN1 PRELIMINARY; PRT; 22 AA.

AC Q9TTN1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Small nuclear ribonucleoprotein B' (Fragment).

GN SNRPB.

OS Monodelphis domestica (Short-tailed grey opossum).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Monodelphis.
OX NCBI_TaxID=13616;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RA  Gray T.A., Martone D.L., Schaldach M.A., McCarrey J.R., Nicholls R.D.;
RT  "Concerted Regulation and Molecular Evolution of the Duplicated SNRPB
RL  Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF134834; AAD54490.1; -.
DR  EMBL; AF134833; AAD54490.1; JOINED.
DR  GO; GO:0030529; C:ribonucleoprotein complex; IEA.
DR  GO; GO:0019013; C:viral nucleocapsid; IEA.
KW  Nucleocapsid; Ribonucleoprotein.
FT  NON_TER 1
SQ  SEQUENCE 22 AA; 2333 MW; 422142330BE11AE9 CRC64;

Query Match 23.3%; Score 30; DB 6; Length 22;
Best Local Similarity 62.5%; Pred. No. 1.7e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 PPPRKMKG 23
DB 4 PPPPGMRG 11

RESULT 7
ID Q9PV93 PRELIMINARY; PRT; 23 AA.
AC Q9PV93;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Small nuclear ribonucleoprotein B' (Fragment).
GN SNRPB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Gray T.A., Martone D.L., Schaldach M.A., McCarrey J.R., Nicholls R.D.;
RT "Concerted Regulation and Molecular Evolution of the Duplicated SNRPB
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF134831; AAD54486.1; -.
DR  GO; GO:0030529; C:ribonucleoprotein complex; IEA.
DR  GO; GO:0019013; C:viral nucleocapsid; IEA.
KW  Nucleocapsid; Ribonucleoprotein.
FT  NON_TER 1
SQ  SEQUENCE 23 AA; 2430 MW; 422142330BE17182 CRC64;

Query Match 23.3%; Score 30; DB 13; Length 23;
Best Local Similarity 62.5%; Pred. No. 1.7e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 16 PPPRKMKG 23
DB 5 PPPPGMRG 12

RESULT 8
ID Q97SX5 PRELIMINARY; PRT; 30 AA.
AC Q97SX5;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SP0188.
GN SP0188.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
```

```

RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC BAA-334 / TIGR4;
RX  MEDLINE=21357209; PubMed=11463916;
RA  Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA  Tettelin H., Nelson K.E., Paulsen I.T., Haft D.H., Dodson R.J.,
RA  Peterson S., Heidelberg J., DeBoy R.T., Nelson W.C., Peterson J.D.,
RA  Durkin A.S., Gwinn M., Kolonay J.F., Nelson M.R., Radune D.,
RA  Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA  Holtzapple E., Knouri H., Wolf A.M., Uterback T.R., Hansen C.L., E.K.,
RA  McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA  Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA  Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT  "Complete genome sequence of a virulent isolate of Streptococcus
RT  pneumoniae".
RL  Science 293:498-506(2001).
DR  EMBL; AE007334; AAK74369.1; -.
DR  PIR; H95021; H95021.
DR  TIGR; SP0188; -.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 30 AA; 3605 MW; EEF84B81A829E11F CRC64;

Query Match 23.3%; Score 30; DB 16; Length 30;
Best Local Similarity 42.1%; Pred. No. 2.2e+03;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 RLKYAQEELSNEVLPPPRK 20
DB 3 RKYENDEKSOKLKIGRK 21

RESULT 9
ID Q63337 PRELIMINARY; PRT; 26 AA.
AC Q63337;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Metabotropic glutamate receptor subtype 7b (Fragment).
GN MGLUR7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC  STRAIN=Sprague Dawley; TISSUE=Brain cortex;
RX  MEDLINE=99092501; PubMed=9875342;
RA  Corti C., Rimland J.M., Ferraguti F.;
RT  "Cloning and characterization of alternative mRNA forms for the rat
RT  metabotropic glutamate receptors mGluR7 and mGluR8."
RL  Eur. J. Neurosci 10:3629-3641(1998).
DR  EMBL; X96790; CAA65584.1; -.
DR  GO; GO:0004872; F:receptor activity; IEA.
KW  Receptor.
FT  NON_TER 1
SQ  SEQUENCE 26 AA; 2940 MW; B4DE50D491PE4EF4 CRC64;

Query Match 22.5%; Score 29; DB 11; Length 26;
Best Local Similarity 55.6%; Pred. No. 2.7e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 NEVLPPPRK 20
DB 3 NNCIPPRK 11

RESULT 10
ID Q9S8V4 PRELIMINARY; PRT; 27 AA.
AC Q9S8V4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cyclic nucleotide binding phosphatase (Fragment).
```

OS Solanum tuberosum (potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamnids; Solanales; Solanaceae; Solanum.
NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RX MEDLINE=93003376; PubMed=1382614;
RA Polya G.M., Wettenthal R.E.;
RT "Rapid purification and N-terminal sequencing of a potato tuber cyclic
nucleotide binding phosphatase.";
RL Biochim. Biophys. Acta 1159:179-184 (1992).
DR PIR; S28940; S28940.
FT NON TER 1 27
FT NON TER 27 27
SQ SEQUENCE 27 AA; 3064 MW; D5B2B30F6EFCF634 CRC64;

Query Match 22.5%; Score 29; DB 10; Length 27;
Best Local Similarity 41.2%; Pred. No. 2.8e+03;
Matches 7; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 RLKVAQEELSNEVLPPP 18
DB 10 RLAVETNNLQNSKLVP 26

RESULT 11
Q8K362 PRELIMINARY; PRT; 29 AA.
ID Q8K362
AC Q8K362; (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028250; AAH28250.1; -.
KW Hypothetical protein.
FT NON TER 1 1
FT NON TER 1 1
SQ SEQUENCE 29 AA; 3309 MW; C8B76A71D0F30ED6 CRC64;

Query Match 22.5%; Score 29; DB 11; Length 29;
Best Local Similarity 50.0%; Pred. No. 3.1e+03;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 AQEELSNEVLPP 17
DB 3 SEELVSLDPP 14

RESULT 12
P82439 PRELIMINARY; PRT; 15 AA.
ID P82439
AC P82439
DT 01-JUN-2000 (TrEMBLrel. 14, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 200 kDa cell wall protein (Fragment).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamnids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. PETIT HAVANA;
RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,

RA Wojtaszek P., Bolwell G.P.;
RT "Proteomic study of secondary cell wall proteins from transformed
RL tobacco culture.";
RL Planta 0:0-0(2000).
CC -1- SUBCELLULAR LOCATION: CELL WALL.
CC -1- TISSUE SPECIFICITY: XYLEM.
DR GO; GO:0005618; C:cell wall; IEA.
KW Cell wall; Hydroxylation.
FT MOD_RES 6 6 HYDROXYLATION.
FT NON TER 15 15
SQ SEQUENCE 15 AA; 1870 MW; 3E1E05A20A3C5681 CRC64;

Query Match 21.7%; Score 28; DB 10; Length 15;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 PPRK 20
DB 8 PPRK 12

RESULT 13
Q9S1C0 PRELIMINARY; PRT; 21 AA.
ID Q9S1C0
AC Q9S1C0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Transposase (Fragment).
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=20002534; PubMed=10531208;
RA Sawada K., Koikeguchi S., Hongyo H., Sawada S., Miyamoto M., Maeda H.,
RA Nishimura F., Takashiba S., Murayama Y.;
RT "Identification by subtractive hybridization of a novel insertion
sequence specific for virulent strains of Porphyromonas gingivalis.";
RL Infect. Immun. 67:5621-5625(1999).
DR EMBL; AB011547; BAA83478.1; -.
FT NON TER 1 1
FT NON TER 1 1
SQ SEQUENCE 21 AA; 2543 MW; 08D3F838FB3E6AF CRC64;

Query Match 21.7%; Score 28; DB 2; Length 21;
Best Local Similarity 37.5%; Pred. No. 3.1e+03;
Matches 6; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 10 LSNEVLPPPRKMKGLF 25
DB 6 LNEPFTPPPEPNOLF 21

RESULT 14
Q9UCC3 PRELIMINARY; PRT; 21 AA.
ID Q9UCC3
AC Q9UCC3; (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE 34 kDa heparin-RELEASABLE protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=94059921; PubMed=8241100;
RA Novotny W.F., Maffi T., Mehta R.L., Milner P.G.;
RT "Identification of novel heparin-releasable proteins, as well as the
cytokines minkine and pleiotrophin, in human postheparin plasma.";

RL Arterioscler. Thromb. 13:1798-1805(1993).
SQ SEQUENCE 21 AA; 2185 MW; 49F4E5706F4B278 CRC64;

Query Match 21.7%; Score 28; DB 4; Length 21;
Best Local Similarity 66.7%; Pred. No. 3.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 16 PPRKMKGL 24
Db 7 PPEVKWVGL 15

RESULT 15

Q8Q7B9 PRELIMINARY; PRT; 26 AA.
AC Q8Q7B9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Tax protein (Fragment).
OS Simian T-lymphotropic virus 1.
OC Viruses; Retroviral viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=33747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAG-DJA.853;
RX MEDLINE=21571643; PubMed=11714973;
RA Nerrienet E., Meertens L., Kfutwah A., Foupouapouognigni Y.,
RA Gessain A.;
RT "Molecular epidemiology of simian T-lymphotropic virus (STLV) in wild-
RT caught monkeys and apes from Cameroon: a new STLV-1, related to human
RT T-lymphotropic virus subtype F, in a Cercopithecus agilis.";
RL J. Gen. Virol. 82:2973-2977(2001).
DR EMBL; AF384872; AAM18645.1; -.
FT NON_TER 1
SQ SEQUENCE 26 AA; 3004 MW; 5C2FB26ADA2DF8EB CRC64;

Query Match 21.7%; Score 28; DB 15; Length 26;
Best Local Similarity 50.0%; Pred. No. 3.9e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 QEELSNEVLPPP 18
Db 5 EPQLSPGLEPP 16

Search completed: July 15, 2004, 14:51:04
Job time : 33.1667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:36:31 ; Search time 50,1667 Seconds
(without alignments)
140.804 Million cell updates/sec

Title: US-09-171-432A-48

Perfect score: 129

Sequence: 1 WLNPKINLADRMGLSGVQRIKEQ 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 581562

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	25	2 AAW42969	Aaw42969 Immunogen
2	129	100.0	25	4 AAB69448	Aab69448 Synthetic
3	96	74.4	20	2 AAW42929	Aaw42929 Immunogen
4	96	74.4	20	4 AAB69446	Aab69446 Synthetic
5	39	30.2	11	3 AAY76318	Aay76318 Human sec
6	39	30.2	11	3 AAY76317	Aay76317 Fragment
7	39	30.2	11	3 AAY76321	Aay76321 Human sec
8	39	30.2	11	7 ADE11971	Ade11971 Human sec
9	39	30.2	27	3 AAY76319	Aay76319 Human sec
10	39	30.2	27	3 AAY76315	Aay76315 Human sec
11	39	30.2	27	7 ADE11969	Ade11969 Human sec
12	35	27.1	14	5 ABG96117	Abg96117 Cysteine-
13	35	27.1	21	4 AAG68117	Aag68117 Human int
14	35	27.1	25	2 AAR88164	Aar88164 Leader se
15	34	26.4	12	3 AAY76316	Aay76316 Fragment
16	34	26.4	12	3 AAY76320	Aay76320 Human sec
17	34	26.4	12	7 ADE11970	Ade11970 Human sec
18	34	26.4	24	3 AAY56845	Aay56845 Apple LRP
19	33	25.6	27	2 AAY07977	Aay07977 Human sec
20	33	25.6	28	2 AAY45329	Aay45329 Human sec
21	33	25.6	29	4 AAM21788	Aam21788 Peptide #
22	33	25.6	29	4 ABB44159	Abb44159 Peptide #
23	33	25.6	29	4 AAM38113	Aam38113 Peptide #
24	33	25.6	29	4 ABB27040	Abb27040 Protein #
25	33	25.6	29	4 AAM77893	Aam77893 Human bon

ALIGNMENTS

RESULT 1

AAW42969
ID AAW42969 standard; peptide; 25 AA.
XX AC AAW42969;
XX DT 28-APR-1998 (first entry)
XX DE Immunogenic Hepatitis A virus peptide YK-1757.
XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.
XX OS Synthetic.
XX OS Hepatitis A virus.
XX PN WO9740147-A1.
XX PD 30-OCT-1997.
XX PF 18-APR-1997; 97WO-US006891.
XX PR 19-APR-1996; 96US-0015644P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX FI Fields HA, Khudyakov YE;
XX DR WPI; 1997-535831/49.
XX PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
XX response to HAV in a mammal or to detect the presence of antibodies
XX against HAV in a mammal.
XX PS Claim 18; Page 112; 140pp; English.
XX CC The present immunogenic peptide corresponds to an immunogenic epitope of
XX the Hepatitis A virus (HAV). The peptide is substantially similar to a
XX portion of the amino acid sequence of the P2A protein of HAV
XX corresponding to amino acids 792-980. Compositions containing the peptide
XX can be used to induce an immune response to HAV in a mammal. The peptide
XX can also be used to detect the presence of antibodies against HAV in
XX mammalian serum. The peptide can also be used to make an antibody against
XX HAV by administering the peptide to a mammal
XX SQ Sequence 25 AA;

Query Match 100.0%; Score 129; DB 2; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.9e-14; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25
 |||||
 Db 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25

RESULT 2

AA69446
 ID AAB69448 standard; peptide; 25 AA.

XX AC AAB69448;

XX 20-APR-2001 (first entry)

XX Synthetic HAV P2A peptide, SEQ ID NO: 48.

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.

OS Synthetic.

XX WO200105824-A2.

XX 25-JAN-2001.

XX 14-JUL-2000; 2000WO-US019267.

XX 15-JUL-1999; 99US-0144412P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 2001-112681/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.

XX Claim 13; Page 99; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC the peptide in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IgM antibody reactivity

XX Sequence 25 AA;

Query Match 100.0%; Score 129; DB 4; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.9e-14; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25
 |||||

Db 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25

RESULT 3

AAW42929

ID AAW42929 standard; peptide; 20 AA.

XX AAW42929;

XX 28-APR-1998 (first entry)

XX Immunogenic Hepatitis A virus peptide YK-1331.

XX Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
 KW antibody.

XX Synthetic.

OS Hepatitis A virus.

XX WO9740147-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-US006891.

XX 19-APR-1996; 96US-0015644P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 1997-535831/49.

XX Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.

XX Claim 18; Page 112; 140pp; English.

XX Peptides AAW42922-30 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the P2A
 CC protein of HAV corresponding to amino acids 792-980. The present peptide
 CC is derived from amino acids 961-980, and has a reactivity of 27.1% with
 CC acute sera. Compositions containing the peptides can be used to induce an
 CC immune response to HAV in a mammal. The peptides can also be used to
 CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal

XX Sequence 20 AA;

Query Match 74.4%; Score 96; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.9e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 KINLADRLMLGLSGVQEIKEQ 25
 |||||

Db 1 KINLADRLMLGLSGVQEIKEQ 20

RESULT 4

AA69446

ID AAB69446 standard; peptide; 20 AA.

XX AAB69446;

XX 20-APR-2001 (first entry)

XX Synthetic HAV P2A peptide, SEQ ID NO: 46.

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.

OS Synthetic.

XX WO200105824-A2.

XX PD 25-JAN-2001.
 XX PF 14-JUL-2000; 2000WO-US019267.
 XX PR 15-JUL-1999; 99US-0144412P.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Fields HA, Khudyakov YE;
 XX PR WPI; 2001-112681/12.
 PT Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 PS Claim 13; Page 97; 130pp; English.
 CC The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IGM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IGM antibody reactivity
 XX Sequence 20 AA;
 SQ Query Match 74.4%; Score 96; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.9e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 KINLADRLGLSGVQEIKEQ 25
 DB 1 KINLADRLGLSGVQEIKEQ 20
 RESULT 5
 AAY76318
 ID AAY76318 standard; protein; 11 AA.
 XX AC AAY76318;
 XX DT 23-MAR-2000 (first entry)
 XX DE Human secreted protein fragment.
 XX KW Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
 XX therapy.
 XX OS Homo sapiens.
 XX FN WO9958660-A1.
 XX PD 18-NOV-1999.
 XX PF 06-MAY-1999; 99WO-US009847.
 XX PR 12-MAY-1998; 98US-0085093P.

PR 12-MAY-1998; 98US-0085094P.
 PR 12-MAY-1998; 98US-0085105P.
 PR 12-MAY-1998; 98US-0085180P.
 PR 18-MAY-1998; 98US-0085906P.
 PR 18-MAY-1998; 98US-0085920P.
 PR 18-MAY-1998; 98US-0085921P.
 PR 18-MAY-1998; 98US-0085922P.
 PR 18-MAY-1998; 98US-0085923P.
 PR 18-MAY-1998; 98US-0085924P.
 PR 18-MAY-1998; 98US-0085925P.
 PR 18-MAY-1998; 98US-0085927P.
 PR 18-MAY-1998; 98US-0085928P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR, Lafleur DW;
 PI Endress GA, Ebner R;
 XX WPI; 2000-062296/05.
 XX New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders.
 XX Disclosure; Page 444; 475pp; English.
 CC AA265250 to AA265350 represent 97 isolated human secreted protein genes.
 CC AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
 CC genes. This sequence represents a fragment of one of the human secreted
 CC proteins. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions, e.g.
 CC by protein or gene therapy. Also pathological conditions can be diagnosed
 CC by determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the new genes. Specific uses are
 CC described for each of the 97 genes, based on which tissues they are most
 CC highly expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumours, developmental abnormalities and foetal
 CC deficiencies, blood disorders, diseases of the immune system, autoimmune
 CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
 CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
 CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
 CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
 CC also useful for identifying their binding partners. The sequences shown
 CC in AAY76224 to AAY76424 represent fragments of the secreted proteins
 XX Sequence 11 AA;
 SQ Query Match 30.2%; Score 39; DB 3; Length 11;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 WLNPKKINL 9
 DB 3 WLNPAKLSL 11
 RESULT 6
 AAY76317
 ID AAY76317 standard; protein; 11 AA.
 XX AC AAY76317;
 XX DT 23-MAR-2000 (first entry)
 XX DE Fragment of human secreted protein encoded by gene 30.
 XX KW Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;

```

XX KW therapy.
XX OS Homo sapiens.
XX PN WO9958660-A1.
XX XX
XX PD 18-NOV-1999.
XX PF
XX PP
XX XX 06-MAY-1999; 99WO-US009847.
XX PR 12-MAY-1998; 98US-0085093P.
XX PR 12-MAY-1998; 98US-0085094P.
XX PR 12-MAY-1998; 98US-0085105P.
XX PR 12-MAY-1998; 98US-0085180P.
XX PR 18-MAY-1998; 98US-0085906P.
XX PR 18-MAY-1998; 98US-0085920P.
XX PR 18-MAY-1998; 98US-0085921P.
XX PR 18-MAY-1998; 98US-0085922P.
XX PR 18-MAY-1998; 98US-0085924P.
XX PR 18-MAY-1998; 98US-0085925P.
XX PR 18-MAY-1998; 98US-0085927P.
XX PR 18-MAY-1998; 98US-0085928P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX XX
XX PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
XX PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR, Lafleur DW;
XX PI Endress GA, Ebner R;
XX DR WPT; 2000-062296/05.
XX XX
XX PT New isolated human genes and the secreted polypeptides they encode,
XX PT useful for diagnosis and treatment of e.g. cancers, neurological
XX PT disorders, immune diseases, inflammation or blood disorders.
XX PS Disclosure; Page 444; 475pp; English.
XX XX
XX CC AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
XX CC AAZ6124 to AAZ76223 are the secreted proteins encoded by the 97 human
XX CC genes. This sequence represents a fragment of one of the human secreted
XX CC proteins. The genes and their corresponding secreted polypeptides are
XX CC useful for preventing, treating or ameliorating medical conditions, e.g.
XX CC by protein or gene therapy. Also pathological conditions can be diagnosed
XX CC by determining the amount of the new polypeptides in a sample or by
XX CC determining the presence of mutations in the new genes. Specific uses are
XX CC described for each of the 97 genes, based on which tissues they are most
XX CC highly expressed in, and include developing products for the diagnosis or
XX CC treatment of cancer, tumours, developmental abnormalities and foetal
XX CC deficiencies, blood disorders, diseases of the immune system, autoimmune
XX CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
XX CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
XX CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
XX CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
XX CC also useful for identifying their binding partners. The sequences shown
XX CC in AAZ76224 to AAZ76424 represent fragments of the secreted proteins
XX XX
XX SQ Sequence 11 AA;
XX
Query Match 30.2%; Score 39; DB 3; Length 11;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 WLNPKKINL 9
Db ||||| :|:|
3 WLNPAKLSL 11
AC
RESULT 7
AAZ76321
ID AAZ76321 standard; protein; 11 AA.
XX
XX AC AAZ76321;

```

```

XX DT 23-MAR-2000 (first entry)
XX DE Human secreted protein fragment.
XX XX
XX KW Human; secreted protein; cancer; tumour; developmental abnormality;
XX KW foetal deficiency; blood disorder; immune system disorder; inflammation;
XX KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
XX KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
XX KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
XX KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
XX KW therapy.
XX OS Homo sapiens.
XX PN WO9958660-A1.
XX XX
XX PD 18-NOV-1999.
XX PF 06-MAY-1999; 99WO-US009847.
XX PR 12-MAY-1998; 98US-0085093P.
XX PR 12-MAY-1998; 98US-0085094P.
XX PR 12-MAY-1998; 98US-0085105P.
XX PR 12-MAY-1998; 98US-0085180P.
XX PR 18-MAY-1998; 98US-0085906P.
XX PR 18-MAY-1998; 98US-0085920P.
XX PR 18-MAY-1998; 98US-0085921P.
XX PR 18-MAY-1998; 98US-0085922P.
XX PR 18-MAY-1998; 98US-0085924P.
XX PR 18-MAY-1998; 98US-0085925P.
XX PR 18-MAY-1998; 98US-0085927P.
XX PR 18-MAY-1998; 98US-0085928P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX XX
XX PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
XX PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR, Lafleur DW;
XX PI Endress GA, Ebner R;
XX DR WPT; 2000-062296/05.
XX XX
XX PT New isolated human genes and the secreted polypeptides they encode,
XX PT useful for diagnosis and treatment of e.g. cancers, neurological
XX PT disorders, immune diseases, inflammation or blood disorders.
XX PS Disclosure; Page 445; 475pp; English.
XX XX
XX CC AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
XX CC AAZ76124 to AAZ76223 are the secreted proteins encoded by the 97 human
XX CC genes. This sequence represents a fragment of one of the human secreted
XX CC proteins. The genes and their corresponding secreted polypeptides are
XX CC useful for preventing, treating or ameliorating medical conditions, e.g.
XX CC by protein or gene therapy. Also pathological conditions can be diagnosed
XX CC by determining the amount of the new polypeptides in a sample or by
XX CC determining the presence of mutations in the new genes. Specific uses are
XX CC described for each of the 97 genes, based on which tissues they are most
XX CC highly expressed in, and include developing products for the diagnosis or
XX CC treatment of cancer, tumours, developmental abnormalities and foetal
XX CC deficiencies, blood disorders, diseases of the immune system, autoimmune
XX CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
XX CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
XX CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
XX CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
XX CC also useful for identifying their binding partners. The sequences shown
XX CC in AAZ76224 to AAZ76424 represent fragments of the secreted proteins
XX XX
XX SQ Sequence 11 AA;
XX
Query Match 30.2%; Score 39; DB 3; Length 11;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 WLNPKKINL 9
Db 3 WLNPAKLSL 11
RESULT 8
ADE11971
ID ADE11971 standard; protein; 11 AA.
AC ADE11971;
XX
XX 29-JAN-2004 (first entry)
DT
XX
DE Human secreted polypeptide #225.
XX
KW Secreted protein; cancer; liver disorder; hepatitis; neural disorder;
KW Alzheimer's disease; human.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX US2003100051-A1.
PN
XX
PD
XX
XX 29-MAY-2003.
XX
PF 10-SEP-2001; 2001US-00948783.
XX
PR 12-MAY-1998; 98US-0085093P.
PR 12-MAY-1998; 98US-0085094P.
PR 12-MAY-1998; 98US-0085105P.
PR 12-MAY-1998; 98US-0085180P.
PR 18-MAY-1998; 98US-0085906P.
PR 18-MAY-1998; 98US-0085920P.
PR 18-MAY-1998; 98US-0085921P.
PR 18-MAY-1998; 98US-0085922P.
PR 18-MAY-1998; 98US-0085923P.
PR 18-MAY-1998; 98US-0085924P.
PR 18-MAY-1998; 98US-0085925P.
PR 18-MAY-1998; 98US-0085926P.
PR 18-MAY-1998; 98US-0085927P.
PR 18-MAY-1998; 98US-0085928P.
PR 06-MAY-1999; 99WO-US009847.
PR 10-NOV-1999; 99US-00437658.
PR 11-SEP-2000; 2000US-0231846P.
PR 28-JUN-2001; 2001US-00892877.
XX
XX (RUBE/) RUBEN S M.
PA (FLOR/) FLORENCE K A.
PA (NIJ/) NI J.
PA (ROSE/) ROSEN C A.
PA (CART/) CARTER K C.
PA (MOOR/) MOORE P A.
PA (OLSE/) OLSEN H S.
PA (SHIY/) SHI Y.
PA (YOUN/) YOUNG P E.
PA (WEIY/) WEI Y.
PA (BREW/) BREWER L A.
PA (SOPP/) SOPPET D R.
PA (LAPL/) LAFLEUR D W.
PA (ENDR/) ENDRESS G A.
PA (EBNE/) EBNER R.
PA (BIRS/) BIRSE C E.
XX
XX Ruben SM, Florence KA, Ni J, Rosen CA, Carter KC, Moore PA;
PI Olsen HS, Shi Y, Young PE, Wei Y, Brewer LA, Soppet DR, Lafleur DW;
PI Endress GA, Ebner R, Birse CE;
XX
XX WPI; 2003-801210/75.
XX
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX

PS Claim 11; SEQ ID NO 343; 453pp; English.
XX
CC The invention relates to human secreted polypeptides and the
CC polynucleotides encoding them. The sequences are useful for preparing
CC medicaments for preventing, treating or ameliorating medical conditions
CC e.g., cancer, liver disorders such as hepatitis or neural disorders such
CC as Alzheimer's disease. This sequence represents a human secreted
CC polypeptide of the invention.
XX

SQ Sequence 11 AA;

Query Match 30.2%; Score 39; DB 7; Length 11;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLNPKKINL 9
Db 3 WLNPAKLSL 11

RESULT 9

AAY76319
ID AAY76319 standard; protein; 27 AA.

XX AAY76319;

DT 23-MAR-2000 (first entry)

DE Human secreted protein fragment.

XX Human; secreted protein; cancer; tumour; developmental abnormality;
KW foetal deficiency; blood disorder; immune system disorder; inflammation;
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KW therapy.

XX Homo sapiens.

XX WO9958660-A1.

PD 18-NOV-1999.

PF 06-MAY-1999; 99WO-US009847.

PR 12-MAY-1998; 98US-0085093P.

PR 12-MAY-1998; 98US-0085094P.

PR 12-MAY-1998; 98US-0085105P.

PR 12-MAY-1998; 98US-0085180P.

PR 18-MAY-1998; 98US-0085906P.

PR 18-MAY-1998; 98US-0085920P.

PR 18-MAY-1998; 98US-0085921P.

PR 18-MAY-1998; 98US-0085922P.

PR 18-MAY-1998; 98US-0085923P.

PR 18-MAY-1998; 98US-0085924P.

PR 18-MAY-1998; 98US-0085925P.

PR 18-MAY-1998; 98US-0085927P.

PR 18-MAY-1998; 98US-0085928P.

PA (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;

PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR, Lafleur DW;

PI Endress GA, Ebner R;

XX WPI; 2000-062296/05.

XX New isolated human genes and the secreted polypeptides they encode,

PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.

PS Disclosure; Page 444; 475pp; English.

XX AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
 CC AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
 CC genes. This sequence represents a fragment of one of the human secreted
 CC proteins. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions, e.g.
 CC by protein or gene therapy. Also pathological conditions can be diagnosed
 CC by determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the new genes. Specific uses are
 CC described for each of the 97 genes, based on which tissues they are most
 CC highly expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumours, developmental abnormalities and foetal
 CC deficiencies, blood disorders, diseases of the immune system, autoimmune
 CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
 CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
 CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
 CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
 CC also useful for identifying their binding partners. The sequences shown
 CC in AAY76224 to AAY76424 represent fragments of the secreted proteins
 XX
 XX Sequence 27 AA;

Query Match 30.2%; Score 39; DB 3; Length 27;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLNPCKINL 9
 ||||| : : :
 Db 17 WLNPAKLSL 25

RESULT 10
 AAY76315
 ID AAY76315 standard; protein; 27 AA.

AC AAY76315;
 DT 23-MAR-2000 (first entry)
 DE Fragment of human secreted protein encoded by gene 30.

XX Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
 KW therapy.

XX Homo sapiens.

XX WO9958660-A1.

XX 18-NOV-1999.

XX 06-MAY-1999; 99WO-US009847.

XX 12-MAY-1998; 98US-0085093P.
 PR 12-MAY-1998; 98US-0085094P.
 PR 12-MAY-1998; 98US-0085105P.
 PR 12-MAY-1998; 98US-0085180P.
 PR 18-MAY-1998; 98US-0085906P.
 PR 18-MAY-1998; 98US-0085920P.
 PR 18-MAY-1998; 98US-0085921P.
 PR 18-MAY-1998; 98US-0085922P.
 PR 18-MAY-1998; 98US-0085923P.
 PR 18-MAY-1998; 98US-0085924P.
 PR 18-MAY-1998; 98US-0085925P.
 PR 18-MAY-1998; 98US-0085927P.
 PR 18-MAY-1998; 98US-0085928P.

(HUMA-) HUMAN GENOME SCI INC.

XX

PI Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 PI Olsen HS, Shi Y, Young PE, Wei F, Brewer LA, Soppet DR, Lafleur DW;
 PI Endress GA, Ener R;
 XX WPI; 2000-062296/05.
 XX
 XX New isolated human genes and the secreted polypeptides they encode,
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders.
 XX
 XX Disclosure; Page 444; 475pp; English.

XX AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
 CC AAY76124 to AAY76223 are the secreted proteins encoded by the 97 human
 CC genes. This sequence represents a fragment of one of the human secreted
 CC proteins. The genes and their corresponding secreted polypeptides are
 CC useful for preventing, treating or ameliorating medical conditions, e.g.
 CC by protein or gene therapy. Also pathological conditions can be diagnosed
 CC by determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the new genes. Specific uses are
 CC described for each of the 97 genes, based on which tissues they are most
 CC highly expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumours, developmental abnormalities and foetal
 CC deficiencies, blood disorders, diseases of the immune system, autoimmune
 CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
 CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
 CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
 CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
 CC also useful for identifying their binding partners. The sequences shown
 CC in AAY76224 to AAY76424 represent fragments of the secreted proteins
 XX
 XX Sequence 27 AA;

Query Match 30.2%; Score 39; DB 3; Length 27;
 Best Local Similarity 66.7%; Pred. No. 40;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLNPCKINL 9
 ||||| : : :
 Db 17 WLNPAKLSL 25

RESULT 11
 ADEL1969
 ID ADEL1969 standard; protein; 27 AA.

AC ADEL1969;

XX 29-JAN-2004 (first entry)

XX Human secreted polypeptide #223.

XX Secreted protein; cancer; liver disorder; hepatitis; neural disorder;
 KW Alzheimer's disease; human.

XX Synthetic.

XX Homo sapiens.

XX US2003100051-A1.

XX 29-MAY-2003.

XX 10-SEP-2001; 2001US-00948783.

XX 12-MAY-1998; 98US-0085093P.

PR 12-MAY-1998; 98US-0085094P.

PR 12-MAY-1998; 98US-0085105P.

PR 12-MAY-1998; 98US-0085180P.

PR 18-MAY-1998; 98US-0085906P.

PR 18-MAY-1998; 98US-0085920P.

PR 18-MAY-1998; 98US-0085921P.

PR 18-MAY-1998; 98US-0085922P.

PR 18-MAY-1998; 98US-0085923P.

PR 18-MAY-1998; 98US-0085924P.
PR 18-MAY-1998; 98US-0085925P.
PR 18-MAY-1998; 98US-0085927P.
PR 18-MAY-1998; 98US-0085928P.
PR 06-MAY-1999; 99WO-US009847.
PR 10-NOV-1999; 99US-00437658.
PR 11-SEP-2000; 2000US-0231846P.
PR 28-JUN-2001; 2001US-00892877.
XX (RUBE/) RUBEN S. M.
PA (FLOK/) FLORENCE K. A.
PA (NIJU/) NI J.
PA (ROSE/) ROSEN C. A.
PA (CART/) CARTER K. C.
PA (MOOR/) MOORE P. A.
PA (OLSE/) OLSEN H. S.
PA (SHIY/) SHI Y.
PA (YOUN/) YOUNG P. E.
PA (WEIY/) WEI Y.
PA (BREW/) BREWER L. A.
PA (SOPP/) SOPPET D. R.
PA (LAEF/) LAFLAUR D. W.
PA (ENDR/) ENDRESS G. A.
PA (EBNE/) EBNER R.
PA (BIRS/) BIRSE C. E.
XX Ruben SM, Florence KA, Ni J, Rosen CA, Carter KC, Moore PA;
PI Olsen HS, Shi Y, Young PE, Wei Y, Brewer LA, Soppet DR, Laflaur DW;
PI Endress GA, Ebner R, Birse CE;
XX WPI; 2003-801210/75.
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX Claim 11; SEQ ID NO 341; 453pp; English.
XX The invention relates to human secreted polypeptides and the
CC polynucleotides encoding them. The sequences are useful for preparing
CC medicaments for preventing, treating or ameliorating medical conditions
CC e.g., cancer, liver disorders such as hepatitis or neural disorders such
CC as Alzheimer's disease. This sequence represents a human secreted
CC polypeptide of the invention.
XX Sequence 27 AA;
Query Match 30.2%; Score 39; DB 7; Length 27;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 WLNPKKINL 9
Db 17 WLNFAKUSL 25
RESULT 12
ABG96117
ID ABG96117 standard; peptide; 14 AA.
XX AC ABG96117;
XX DT 11-DEC-2002 (first entry)
XX Cysteine-containing peptide isolated by the invention method, #90.
DE Rabbit; bovine; analytical reagent; trifunctional; peptide mixture;
KW enrichment; immobilisation site; cleavage site; link; epitope tag;
KW protease; cysteine-containing; perturbed cell; mass spectrometry;
KW peptide tag; BSA; bovine serum albumin; PEPTag; APEPTag; IPEPTag;
KW affinity peptide encoded tag; immobilised peptide encoded tag; chicken;
KW beta-lactoglobulin; GAPDH; glyceraldehyde-3-phosphate dehydrogenase;
KW a-lactalbumin; ovalbumin; yeast.

XX Saccharomycetes cerevisiae.
OS WO200259144-A2.
XX 01-AUG-2002.
XX 25-JAN-2002; 2002WO-US002487.
XX 26-JAN-2001; 2001US-0264576P.
PR 13-JUL-2001; 2001US-0305232P.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Haynes P, Wei J, Yates J, Andon N;
XX WPI; 2002-599760/64.
XX Novel trifunctional synthetic reagents for labeling peptides at specific
PT amino acid residue and selectively enriching only those peptides
PT containing labeled amino acid, useful for proteomic analysis.
XX Disclosure; Page 48; 79pp; English.
XX The invention discloses analytical reagents (e.g. trifunctional synthetic
CC reagents) which can be used for reducing the complexity of peptide
CC mixtures. The method labels peptides at a specific amino acid residue and
CC then selectively enriches only those peptides containing the labelled
CC amino acid. The compound have the formula of immobilisation site-cleavage
CC site-link. The immobilisation site is chosen from an epitope tag, a
CC linker to a solid surface, a metal chelating site, a magnetic site and a
CC specific oligonucleotide sequence, or their combination, the cleavage
CC site is chosen from a protease cleavage site, a photocleavable linker, a
CC restriction enzyme cleavage site, a chemical cleavage site and a thermal
CC cleavage site, or their combination and the link is chosen from an amino
CC acid reactive site and a mass variance site, or their combination. The
CC compounds are useful for simultaneously identifying and determining the
CC levels of expression of cysteine-containing proteins in normal and
CC perturbed cells. The advantage is that these reagents allow rapid and
CC quantitative analysis of proteins or protein function in mixtures of
CC proteins. By preparing the reagent in two forms with detectably different
CC masses, accurate relative quantification of peptide amounts using mass
CC spectrometry, can be achieved. The sequences given in ABG95935-ABG96244
CC are examples of the peptide tags used to isolate cysteine-containing
CC proteins, the target sequences tested and the peptides isolated using the
CC peptide tags
XX Sequence 14 AA;
Query Match 27.1%; Score 35; DB 5; Length 14;
Best Local Similarity 45.5%; Pred. No. 87;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 14 LGLSGVQBIKE 24
Db 4 LGITGVQVRE 14
RESULT 13
AAG68117
ID AAG68117 standard; peptide; 21 AA.
XX AC AAG68117;
XX 22-JAN-2002 (first entry)
XX Human interleukin 1 family protein related peptide SEQ ID NO:12.
DE Human; interleukin 1; IL-1; growth factor; Tango-77; diagnosis;
KW identification.
XX Homo sapiens.
OS

FH Key Location/Qualifiers
 FT Modified-site 21
 FT /label= amidated
 XX
 XX JP2001231578-A.
 XX
 XX 28-AUG-2001.
 PD
 XX 07-DEC-2000; 2000JP-00372864.
 PF
 XX 09-DEC-1999; 99JP-00349780.
 PR
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX WPI; 2001-609968/70.
 DR
 XX
 XX An IL-1 family protein, used for the development of diagnostic and
 PT treatment agents.
 PT
 XX Example 7; Page 24-25; 38pp; Japanese.
 PS
 XX The present invention describes a human interleukin 1 (IL-1) family
 CC protein having a combining affinity to a receptor of a protein of human
 CC IL-1 family higher than Tango-77. The protein is useful for the
 CC development of diagnostic, treating and/or preventive agents for various
 CC diseases. The present sequence represents a human IL-1 family protein
 CC related peptide, which is used in an example from the present invention
 CC
 XX Sequence 21 AA;
 SQ
 Query Match 27.1%; Score 35; DB 4; Length 21;
 Best Local Similarity 54.5%; Pred. No. 1.4e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 LNPCKINLADR 12
 Db ||||| :|||
 7 LNPCKFSIHQ 17
 RESULT 14
 AA88164
 ID AAR88164 standard; peptide; 25 AA.
 AC AAR88164;
 XX
 XX 23-JUL-1996 (first entry)
 DT
 XX Leader sequence for expression of heterologous protein in yeast.
 DE
 XX Leader sequence; aprotinin; insulin; Saccharomyces cerevisiae;
 KW expression; secretion; cassette; alpha-factor; mouse salivary amylase;
 KW carboxypeptidase; yeast; factor 10a; aspartic protease 3; BARI;
 KW KEX2 protease; tissue factor pathway inhibitor; heterologous protein.
 XX
 XX Synthetic.
 OS
 XX WO9534666-A1.
 PN
 XX 21-DEC-1995.
 PD
 XX 16-JUN-1995; 95WO-DK000249.
 XX
 XX 16-JUN-1994; 94DK-00000705.
 PR
 XX 29-JUL-1994; 94US-00282852.
 PR
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Kjeldsen TB, Vad K;
 XX
 XX WPI; 1996-049693/05.
 DR
 XX Expression cassette for yeast contg. synthetic leader sequence -
 PT providing high yields of secreted polypeptide encoded by the cassette,
 PT

PT also related vectors and transformed yeast cells.
 XX Claim 21; Page 66; 85pp; English.
 PS
 XX The present peptide is a leader sequence for the expression of a
 CC heterologous protein (HP) in yeast, providing high level expression and
 CC secretion. An expression cassette for a HP in yeast, comprises 5'-3', a
 CC promoter (P), sequences encoding a signal peptide (SP), the present
 CC peptide, a processing site (PS) and the HP and an optional terminator
 CC sequence. The P can be any P functional in yeast, e.g. the alpha factor
 CC gene P, and the SP is pref. the alpha-factor, mouse salivary amylase,
 CC carboxypeptidase, yeast aspartic protease 3 or yeast BARI SP. The PS is
 CC LysArg, ArgLys, ArgArg or LysLys, for processing by S. cerevisiae KEX2
 CC protease, or IleGluGlyArg, for processing by factor 10a, and the HP can
 CC be, e.g. aprotinin, tissue factor pathway inhibitor, insulin or its
 CC precursors, etc
 XX
 XX Sequence 25 AA;
 SQ
 Query Match 27.1%; Score 35; DB 2; Length 25;
 Best Local Similarity 43.8%; Pred. No. 1.7e+02;
 Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 3 NPKKINLADRLMLGLSG 18
 Db | :||| :|||
 9 NTTSVNLMDLAVGLPG 24
 RESULT 15
 AAY76316
 ID AAY76316 standard; protein; 12 AA.
 XX
 XX AC AAY76316;
 XX
 XX 23-MAR-2000 (first entry)
 DT
 XX Fragment of human secreted protein encoded by gene 30.
 DE
 XX Human; secreted protein; cancer; tumour; developmental abnormality;
 KW foetal deficiency; blood disorder; immune system disorder; inflammation;
 KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
 KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
 KW therapy.
 XX
 XX Homo sapiens.
 OS
 XX WO9958660-A1.
 PN
 XX 18-NOV-1999.
 PD
 XX 06-MAY-1999; 99WO-US009847.
 PF
 XX 12-MAY-1998; 98US-0085093P.
 PR 12-MAY-1998; 98US-0085094P.
 PR 12-MAY-1998; 98US-0085105P.
 PR 12-MAY-1998; 98US-0085180P.
 PR 18-MAY-1998; 98US-0085906P.
 PR 18-MAY-1998; 98US-0085920P.
 PR 18-MAY-1998; 98US-0085921P.
 PR 18-MAY-1998; 98US-0085922P.
 PR 18-MAY-1998; 98US-0085923P.
 PR 18-MAY-1998; 98US-0085924P.
 PR 18-MAY-1998; 98US-0085925P.
 PR 18-MAY-1998; 98US-0085927P.
 PR 18-MAY-1998; 98US-0085928P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Florence K, Ni J, Rosen CA, Carter KC, Moore PA;
 PI Olsen HS, Shi Y, Young FE, Wei F, Brewer LA, Soppet DR, Lafleur DW;
 PI Endress GA, Ebner R;

```

XX WPI, 2000-062296/05.
DR
XX
XX New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
XX Disclosure; Page 444; 475pp; English.
PS
XX
XX AAZ65250 to AAZ65350 represent 97 isolated human secreted protein genes.
CC
CC AAZ76124 to AAY76223 are the secreted proteins encoded by the 97 human
CC genes. This sequence represents a fragment of one of the human secreted
CC proteins. The genes and their corresponding secreted polypeptides are
CC useful for preventing, treating or ameliorating medical conditions, e.g.
CC by protein or gene therapy. Also pathological conditions can be diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new genes. Specific uses are
CC described for each of the 97 genes, based on which tissues they are most
CC highly expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumours, developmental abnormalities and foetal
CC deficiencies, blood disorders, diseases of the immune system, autoimmune
CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
CC also useful for identifying their binding partners. The sequences shown
CC in AAY76224 to AAY76424 represent fragments of the secreted proteins
XX
XX Sequence 12 AA;
SQ
Query Match 26.4%; Score 34; DB 3; Length 12;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 WLNPKKI 7
| | | | |
Db 6 WLNPAKL 12
Search completed: July 15, 2004, 14:46:55
Job time : 53.1667 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:47:02 ; Search time 15.5 Seconds
(without alignments)
83.268 Million cell updates/sec

Title: US-09-171-432A-48
Perfect score: 129
Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 204689

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap: *
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap: *
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap: *
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap: *
5: /cgn2_6/ptodata/2/iaa/6CTUS COMB.pap: *
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	27.1	25	1	US-08-468-674B-15
2	35	27.1	25	1	US-08-780-571-15
3	35	27.1	25	4	US-09-324-217B-17
4	34	26.4	28	1	US-07-977-630-45
5	32	24.8	20	2	US-08-934-915-23
6	31	24.0	25	1	US-08-248-357C-2
7	30	23.3	28	3	US-09-295-186-3
8	30	23.3	30	3	US-09-045-632-79
9	30	23.3	30	3	US-08-695-301A-42
10	30	23.3	30	3	US-08-695-304C-42
11	30	23.3	30	3	US-08-824-800D-6
12	30	23.3	30	3	US-09-588-751-6
13	30	23.3	30	3	US-09-589-768-42
14	30	23.3	30	3	US-09-594-845-42
15	29.5	22.9	28	2	US-08-488-161-48
16	29.5	22.9	28	3	US-09-273-685-48
17	29.5	22.9	28	5	PCT-US95-11934-48
18	29	22.5	14	1	US-07-854-603-B
19	29	22.5	16	1	US-08-447-925-5
20	29	22.5	17	4	US-09-424-712-21
21	29	22.5	22	3	US-08-940-095-173
22	29	22.5	22	3	US-08-940-093-173
23	29	22.5	22	3	US-08-940-096-173
24	29	22.5	22	3	US-09-465-719-173
25	29	22.5	22	4	US-09-453-605-173
26	29	22.5	22	4	US-09-453-838-173
27	29	22.5	22	4	US-08-940-136-173

28	29	22.5	22	4	US-09-453-841-173	Sequence 173, App
29	29	22.5	22	4	US-09-453-833-173	Sequence 173, App
30	29	22.5	22	4	US-09-453-826-173	Sequence 173, App
31	28	21.7	11	3	US-09-215-966-20	Sequence 20, Appl
32	28	21.7	12	1	US-08-196-989B-6	Sequence 6, Appl
33	28	21.7	12	2	US-08-760-936-6	Sequence 6, Appl
34	28	21.7	12	4	US-09-225-024-6	Sequence 6, Appl
35	28	21.7	19	3	US-08-779-764A-41	Sequence 41, Appl
36	28	21.7	19	4	US-09-563-456-41	Sequence 41, Appl
37	28	21.7	20	4	US-09-664-945-26	Sequence 26, Appl
38	28	21.7	23	1	US-08-056-200-57	Sequence 57, Appl
39	28	21.7	23	1	US-08-485-588-10	Sequence 10, Appl
40	28	21.7	23	1	US-08-484-565-10	Sequence 10, Appl
41	28	21.7	23	2	US-08-480-751-10	Sequence 10, Appl
42	28	21.7	23	2	US-08-800-644-57	Sequence 57, Appl
43	28	21.7	23	2	US-08-943-986-10	Sequence 10, Appl
44	28	21.7	23	3	US-08-353-784-10	Sequence 10, Appl
45	28	21.7	23	3	US-08-484-719B-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-468-674B-15
; Sequence 15, Application US/08468674B
; Patent No. 5639642
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas B
; APPLICANT: Vad, Knud
; TITLE OF INVENTION: SYNTHETIC LEADERS PEPTIDE SEQUENCES
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 56396420 No. 5639642disk of No. 5639642th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,674B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/282,852
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4085.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-468-674B-15

Query Match 27.1%; Score 35; DB 1; Length 25;
Best Local Similarity 43.8%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 3 NPKKINLADRMGLSG 18
Db 9 NTTSVNLMDLAVGLPG 24

RESULT 2
US-08-780-571-15
; Sequence 15, Application US/08780571
; Patent No. 5795746
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas B
; APPLICANT: Vad, Knud
; TITLE OF INVENTION: SYNTHETIC LEADERS PEPTIDE SEQUENCES
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5795746 of No. 5795746disk of No. 5795746th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/780,571
; FILING DATE: 08-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/468,674
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: US 08/282,852
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4085-220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-780-571-15

Query Match 27.1%; Score 35; DB 1; Length 25;
Best Local Similarity 43.8%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 NPKKINLADRLMLGSLG 18
| :||| :|||
Db 9 NTTSVNMLDLAVGLPG 24

RESULT 3
US-09-324-217B-17
; Sequence 17, Application US/09324217B
; Patent No. 6500645
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas Boeglum
; APPLICANT: Balschmidt, Per
; APPLICANT: Petersson, Annette Frost
; APPLICANT: Vad, Knud
; APPLICANT: Brandt, Jakob
; APPLICANT: Havelund, Svend
; TITLE OF INVENTION: N-Terminally Extended Proteins Expressed
; TITLE OF INVENTION: In Yeast
; FILE REFERENCE: 5930-500-US
; CURRENT APPLICATION NUMBER: US/09/324,217B
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/036,758

Query Match 26.4%; Score 34; DB 1; Length 28;
Best Local Similarity 46.2%; Pred. No. 48;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 LADRLMLGSLGVQVE 21
| :||| :||| :

; PRIOR FILING DATE: 1997-01-28
; PRIOR APPLICATION NUMBER: 0712/94
; PRIOR FILING DATE: 1994-06-17
; PRIOR APPLICATION NUMBER: 1449/95
; PRIOR FILING DATE: 1995-12-20
; PRIOR APPLICATION NUMBER: 1482/96
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Variation
; US-09-324-217B-17

Query Match 27.1%; Score 35; DB 4; Length 25;
Best Local Similarity 43.8%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 NPKKINLADRLMLGSLG 18
| :||| :|||
Db 9 NTTSVNMLDLAVGLPG 24

RESULT 4
US-07-977-630-45
; Sequence 45, Application US/07977630
; Patent No. 5583038
; GENERAL INFORMATION:
; APPLICANT: Stover, Charles K.
; TITLE OF INVENTION: BACTERIAL EXPRESSION VECTORS CONTAINING
; TITLE OF INVENTION: DNA ENCODING SECRETION SIGNALS OF LIPOPROTEINS
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
; ADDRESSEE: Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,630
; FILING DATE: No. 5583038member 17, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Herron, Charles J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 469201-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-977-630-45

Query Match 26.4%; Score 34; DB 1; Length 28;
Best Local Similarity 46.2%; Pred. No. 48;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 LADRLMLGSLGVQVE 21
| :||| :||| :

```

; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,357C
; FILING DATE: 24-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
; US-08-248-357C-2
;
Query Match 24.0%; Score 31; DB 1; Length 25;
Best Local Similarity 37.5%; Pred. No. 1.4e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 4 PKINLADRLGLS-----GVQE 21
DB 1 PPKKEIPDRIRGRSRTSGSPGLQE 24

RESULT 7
US-09-295-186-3
; Sequence 3, Application US/09295186B
; Patent No. 6127137
; GENERAL INFORMATION:
; APPLICANT: Hasida, Miyoko
; APPLICANT: Tsutsumi, No. 6127137iko
; APPLICANT: Halkier, Torben
; APPLICANT: Stringer, Mary Ann
; TITLE OF INVENTION: An Acidic Phospholipase, Production, and
; TITLE OF INVENTION: Methods of Using Thereof (As Amended)
; FILE REFERENCE: 4953.204-US
; CURRENT APPLICATION NUMBER: US/09/295,186B
; CURRENT FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 1215/96
; PRIOR FILING DATE: 1996-10-31
; PRIOR APPLICATION NUMBER: PCT/DK97/00490
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Hypozyma sp. CBS 648.91
; FEATURE:
; OTHER INFORMATION: Xaa at position 4 is any amino acid
; US-09-295-186-3
;
Query Match 23.3%; Score 30; DB 3; Length 28;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 INLADRLGLS 18
DB 15 LQAEVINGSL 27

```

RESULT 8
 US-09-045-632-79
 ; Sequence 79, Application US/09045632
 ; Patent No. 6001575
 ; GENERAL INFORMATION:
 ; APPLICANT: HUGANIR, Richard L.
 ; APPLICANT: Dong, Hualing
 ; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
 ; TITLE OF INVENTION: GRIP-RELATED MOLECULES
 ; NUMBER OF SEQUENCES: 105
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 ; STREET: 130 Water Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/045,632
 ; FILING DATE: 19-MAR-1998
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/041,016
 ; FILING DATE: 19-MAR-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Corless, Peter F.
 ; REGISTRATION NUMBER: 33,860
 ; REFERENCE/DOCKET NUMBER: 48147/1699-CIP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-523-3400
 ; TELEFAX: 617-523-6440
 ; INFORMATION FOR SEQ ID NO: 79:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 30 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-045-632-79

Query Match 23.3%; Score 30; DB 3; Length 30;
 Best Local Similarity 23.1%; Pred. No. 2.6e+02;
 Matches 3; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 7 INLADRLGLSGV 19
 Db 17 LQIGDRVMAINGI 29

RESULT 9
 US-08-695-301A-42
 ; Sequence 42, Application US/08695301A
 ; Patent No. 6093400
 ; GENERAL INFORMATION:
 ; APPLICANT: DANIEL H. ZIMMERMAN, PREM S. SARIN
 ; TITLE OF INVENTION: Modified HGP-30 Peptides, Conjugates,
 ; TITLE OF INVENTION: Compositions and Methods of Use
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Law Office of Sherman and Shalloway
 ; STREET: 413 N. Washington Street
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22314
 ; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 1.2 mb storage
 ; MEDIUM TYPE: 3.5 inch, 1.44 mb storage
 ; COMPUTER: Dell System 210; Intel 80 286 Microprocessor
 ; OPERATING SYSTEM: MS DOS 6.22
 ; SOFTWARE: Word Perfect, Version 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/695,301A
 ; FILING DATE: August 9, 1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Richard A. Steinberg
 ; REGISTRATION NUMBER: 26,588
 ; REFERENCE/DOCKET NUMBER: CELL-101
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 549-2282
 ; TELEFAX: (703) 836-0106
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 30 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal fragment
 ; FEATURE:
 ; NAME/KEY:
 ; LOCATION: 85 to 114
 ; OTHER INFORMATION: fragment of p-17 gag protein of
 ; OTHER INFORMATION: HIV, Thailand-NE
 ; US-08-695-301A-42

Query Match 23.3%; Score 30; DB 3; Length 30;
 Best Local Similarity 20.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRLGLSGVQRIKEQ 25
 Db 1 WCVHQRIEVKDTKEALDKIEVQNK 25

RESULT 10
 US-08-695-304C-42
 ; Sequence 42, Application US/08695304C
 ; Patent No. 6103239
 ; GENERAL INFORMATION:
 ; APPLICANT: DANIEL H. ZIMMERMAN, PREM S. SARIN
 ; TITLE OF INVENTION: Modified HGP-30 Heteroconjugates,
 ; TITLE OF INVENTION: Compositions and Methods of Use
 ; NUMBER OF SEQUENCES: 56
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Law Office of Sherman and Shalloway
 ; STREET: 413 N. Washington Street
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22314
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 5.25 inch, 1.2 mb storage
 ; MEDIUM TYPE: 3.5 inch, 1.44 mb storage
 ; COMPUTER: Dell System 210; Intel 80 286 Microprocessor
 ; OPERATING SYSTEM: MS DOS 6.22
 ; SOFTWARE: Word Perfect, Version 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/695,304C
 ; FILING DATE: August 9, 1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Richard A. Steinberg
 ; REGISTRATION NUMBER: 26,588
 ; REFERENCE/DOCKET NUMBER: CELL-102
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 549-2282
 ; TELEFAX: (703) 836-0106
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 30 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
FEATURE:
NAME/KEY:
LOCATION: 85 to 114
OTHER INFORMATION: fragment of p-17 gag protein of HIV; Thailand-NE
US-08-695-304C-42

Query Match 23.3%; Score 30; DB 3; Length 30;
Best Local Similarity 20.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 1 WCVHQRIEVKDTKEALDKIEEVQNK 25

RESULT 11
US-08-824-800D-6
Sequence 6, Application US/08624800D
Patent No. 6111068
GENERAL INFORMATION:
APPLICANT: DANIEL H. ZIMMERMAN AND PREM S. SARIN
TITLE OF INVENTION: HIV P-17 PEPTIDE FRAGMENT, COMPOSITIONS
TITLE OF INVENTION: CONTAINING AND METHODS FOR PRODUCING AND USING SAME
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Office of Sherman and Shalloway
STREET: 413 N. Washington Street
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: Dell System 210; Intel 80 286 Microprocessor
OPERATING SYSTEM: MS DOS 6.2
SOFTWARE: Word Perfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,800D
FILING DATE: March 26, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Richard A. Steinberg
REGISTRATION NUMBER: 26,588
REFERENCE/DOCKET NUMBER: VT-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 549-2282
TELEFAX: (703) 836-0106
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acids
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: HIV-1
FEATURE:
NAME/KEY: Thailand E Consensus sequence
LOCATION: 1-30
IDENTIFICATION METHOD: established consensus sequence
OTHER INFORMATION: Consensus E sequence, residues of p17
OTHER INFORMATION: gag protein at positions 86-115
US-08-824-800D-6

Query Match 23.3%; Score 30; DB 3; Length 30;
Best Local Similarity 20.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 1 WCVHQRIEVKDTKEALDKIEEVQNK 25

RESULT 12
US-09-588-751-6
Sequence 6, Application US/09588751
Patent No. 6258945
GENERAL INFORMATION:
APPLICANT: DANIEL H. ZIMMERMAN AND PREM S. SARIN
TITLE OF INVENTION: HIV P-17 PEPTIDE FRAGMENT, COMPOSITIONS
TITLE OF INVENTION: CONTAINING AND METHODS FOR PRODUCING AND USING SAME
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Office of Sherman and Shalloway
STREET: 413 N. Washington Street
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: Dell System 210; Intel 80 286 Microprocessor
OPERATING SYSTEM: MS DOS 6.2
SOFTWARE: Word Perfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/588,751
FILING DATE: Concurrently herewith
ATTORNEY/AGENT INFORMATION:
NAME: Richard A. Steinberg
REGISTRATION NUMBER: 26,588
REFERENCE/DOCKET NUMBER: VT-8.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 549-2282
TELEFAX: (703) 836-0106
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acids
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: HIV-1
FEATURE:
NAME/KEY: Thailand E Consensus sequence
LOCATION: 1-30
IDENTIFICATION METHOD: established consensus sequence
OTHER INFORMATION: Consensus E sequence, residues of p17
OTHER INFORMATION: gag protein at positions 86-115
US-09-588-751-6

Query Match 23.3%; Score 30; DB 3; Length 30;
Best Local Similarity 20.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 1 WCVHQRIEVKDTKEALDKIEEVQNK 25

RESULT 13
US-09-589-768-42
Sequence 42, Application US/09589768
Patent No. 6268472
GENERAL INFORMATION:
APPLICANT: DANIEL H. ZIMMERMAN, PREM S. SARIN
TITLE OF INVENTION: Modified HGP-30 Peptides, Conjugates,
TITLE OF INVENTION: Compositions and Methods of Use
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:

Query Match 23.3%; Score 30; DB 3; Length 30;
Best Local Similarity 20.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 WLN-----PKKINLADR 12
||| | | |
Db 8 WLNQLGSAALPSKENSQDR 26

Search completed: July 15, 2004, 15:13:24
Job time : 16.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 15:11:48 ; Search time 40.6667 Seconds
(without alignments)
192.148 Million cell updates/sec

Title: US-09-171-432A-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRLGLSGVQEIKEQ 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 288454

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	129	100.0	25	US-09-171-432A-48	Sequence 48, Appl
2	96	74.4	20	US-09-171-432A-46	Sequence 46, Appl
3	39	30.2	11	US-09-892-877-330	Sequence 330, Appl
4	39	30.2	11	US-09-948-783-343	Sequence 343, Appl
5	39	30.2	27	US-09-892-877-328	Sequence 328, Appl
6	39	30.2	27	US-09-948-783-341	Sequence 341, Appl
7	34	26.4	12	US-09-892-877-329	Sequence 329, Appl
8	34	26.4	12	US-09-948-783-342	Sequence 342, Appl
9	34	26.4	30	US-10-437-963-186323	Sequence 186323,
10	33	25.6	27	US-10-195-730-291	Sequence 291, Appl
11	33	25.6	28	US-09-798-889-121	Sequence 121, Appl
12	33	25.6	28	US-10-633-680-121	Sequence 121, Appl
13	33	25.6	29	US-09-864-761-42338	Sequence 42338, A
14	33	25.6	29	US-10-695-499-147	Sequence 147, Appl
15	32	24.8	20	US-10-112-488-53	Sequence 53, Appl

16 32 24.8 20 16 US-10-673-860-40
17 32 24.8 24 14 US-10-271-078-12
18 32 24.8 25 14 US-10-172-088-16
19 32 24.8 25 16 US-10-416-249-149
20 32 24.8 27 9 US-09-864-761-38784
21 31 24.0 9 12 US-09-889-468-34
22 31 24.0 20 14 US-10-087-464-3
23 31 24.0 20 14 US-10-225-567A-1926
24 31 24.0 24 15 US-10-424-233-44
25 31 24.0 28 9 US-09-848-035-10
26 31 24.0 28 9 US-09-815-626-11
27 31 24.0 28 9 US-09-986-224-10
28 31 24.0 28 13 US-10-127-516-12
29 31 24.0 28 13 US-10-027-629-12
30 31 24.0 28 14 US-10-145-586-70
31 31 24.0 28 14 US-10-132-967-12
32 30.5 23.6 26 12 US-10-425-114-6887
33 30 23.3 12 14 US-10-057-789-189
34 30 23.3 12 14 US-10-213-628-189
35 30 23.3 15 9 US-09-829-549A-36
36 30 23.3 17 16 US-10-416-249-322
37 30 23.3 17 16 US-10-273-907A-17
38 30 23.3 20 9 US-09-731-221-71
39 30 23.3 21 14 US-10-206-155-1
40 30 23.3 21 14 US-10-360-836-3
41 30 23.3 25 16 US-10-416-249-151
42 30 23.3 25 16 US-10-416-249-152
43 30 23.3 29 12 US-10-424-599-146881
44 30 23.3 30 12 US-10-296-734-1326
45 30 23.3 30 12 US-10-296-317-17

ALIGNMENTS

RESULT 1

US-09-171-432A-48
; Sequence 48, Application US/09171432A
; Publication NO. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

Fri Jul 16 09:25:23 2004

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..25
OTHER INFORMATION: /label= YK-1757
US-09-171-432A-48
Query Match 100.0%; Score 129; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.6e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25
Db 1 WLNPKKINLADRLMLGLSGVQEIKEQ 25
RESULT 2
US-09-171-432A-46
Sequence 46, Application US/09171432A
Publication No. US20030187184A1
GENERAL INFORMATION:
APPLICANT: Fields, Howard A.
APPLICANT: Khudyakov, Yuri E.
TITLE OF INVENTION: Antigenically Reactive Regions of the
TITLE OF INVENTION: Hepatitis A Virus Polyprotein
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick Stockton LLP
STREET: 3424 Peachtree Road, N.E.
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30326
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09171432A
APPLICATION NUMBER: US/09171432A
FILING DATE: 23-NOV-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 03063-0231US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 949-2400
TELEFAX: (404) 949-2499
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /label= YK-1331
US-09-171-432A-46
Query Match 74.4%; Score 96; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 KINLADRLMLGLSGVQEIKEQ 25
Db 1 KINLADRLMLGLSGVQEIKEQ 20
RESULT 3
US-09-892-877-330
Sequence 330, Application US/09892877
Publication No. US20030077809A1
GENERAL INFORMATION:
APPLICANT: Ruben et. al.
TITLE OF INVENTION: 97 Human secreted proteins
FILE REFERENCE: PZ028P1
CURRENT APPLICATION NUMBER: US/09/892,877
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658
PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-10
NUMBER OF SEQ ID NOS: 461
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 330
LENGTH: 11
TYPE: PRT
ORGANISM: Homo sapiens
US-09-892-877-330
Query Match 30.2%; Score 39; DB 10; Length 11;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 WLNPKKINL 9
Db 3 WLNPAKLSL 11
RESULT 4
US-09-948-783-343
Sequence 343, Application US/09948783
Publication No. US20030100051A1
GENERAL INFORMATION:
APPLICANT: Ruben et. al.
TITLE OF INVENTION: 97 Human secreted proteins
FILE REFERENCE: PZ028P2
CURRENT APPLICATION NUMBER: US/09/948,783
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/231,846
PRIOR FILING DATE: 2000-09-11
PRIOR APPLICATION NUMBER: 09/892,877
PRIOR FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 09/437,658
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: PCT/US99/09847
PRIOR FILING DATE: 1999-05-06
PRIOR APPLICATION NUMBER: 60/085,093
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,094
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,105
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,180
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,927
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,906
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,924
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,922
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,921
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,923
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,925

```
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,928
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,920
; PRIOR FILING DATE: 1998-05-18
; NUMBER OF SEQ ID NOS: 465
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 343
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-948-783-343

Query Match          30.2%; Score 39; DB 10; Length 11;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WLNPKKINL 9
Db 3 WLNPAKLSL 11

RESULT 5
US-09-892-877-328
; Sequence 328, Application US/09892877
; Publication No. US20030077809A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: PZ028P1
; CURRENT APPLICATION NUMBER: US/09/892,877
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 328
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-892-877-328

Query Match          30.2%; Score 39; DB 10; Length 27;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WLNPKKINL 9
Db 17 WLNPAKLSL 25

RESULT 6
US-09-948-783-341
; Sequence 341, Application US/09948783
; Publication No. US20030100051A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: PZ028P2
; CURRENT APPLICATION NUMBER: US/09/948,783
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,846
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: 09/892,877
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/437,658
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/09847
; PRIOR FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: 60/085,093
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,094
; PRIOR FILING DATE: 1998-05-12
```

```
; PRIOR APPLICATION NUMBER: 60/085,105
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,180
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,927
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,906
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,924
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,928
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,920
; PRIOR FILING DATE: 1998-05-18
; NUMBER OF SEQ ID NOS: 465
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 341
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-948-783-341

Query Match          30.2%; Score 39; DB 10; Length 27;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WLNPKKINL 9
Db 17 WLNPAKLSL 25

RESULT 7
US-09-892-877-329
; Sequence 329, Application US/09892877
; Publication No. US20030077809A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: PZ028P1
; CURRENT APPLICATION NUMBER: US/09/892,877
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 329
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-892-877-329

Query Match          26.4%; Score 34; DB 10; Length 12;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WLNPKKI 7
Db 6 WLNPAKL 12

RESULT 8
US-09-948-783-342
; Sequence 342, Application US/09948783
; Publication No. US20030100051A1
; GENERAL INFORMATION:
```

APPLICANT: Ruben et. al.
TITLE OF INVENTION: 97 Human secreted proteins
FILE REFERENCE: P2028P2
CURRENT APPLICATION NUMBER: US/09/948,783
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/231,846
PRIOR FILING DATE: 2000-09-11
PRIOR APPLICATION NUMBER: 09/892,877
PRIOR FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 09/437,658
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: PCT/US99/09847
PRIOR FILING DATE: 1999-05-06
PRIOR APPLICATION NUMBER: 60/085,093
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,094
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,105
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,180
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085,927
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,906
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,924
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,922
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,921
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,923
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,925
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,928
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,920
PRIOR FILING DATE: 1998-05-18
NUMBER OF SEQ ID NOS: 465
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 342
LENGTH: 12
TYPE: PRT
ORGANISM: Homo sapiens
US-09-948-783-342

Query Match 26.4%; Score 34; DB 10; Length 12;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLNPKKI 7
DB 6 WLNPAKL 12

RESULT 9
US-10-437-963-186323
Sequence 186323, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Xihua
APPLICANT: Cao, Jiongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 186323
LENGTH: 30
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_83133C.1.pap
US-10-437-963-186323

Query Match 26.4%; Score 34; DB 16; Length 30;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 NLADRMGLSGV 19
DB 3 NCYDRYVGLTGI 14

RESULT 10
US-10-195-730-291
Sequence 291, Application US/10195730
Publication No. US20030144492A1
GENERAL INFORMATION:
APPLICANT: Rosen et. al.
TITLE OF INVENTION: 101 Human Secreted Proteins
FILE REFERENCE: P2017P1
CURRENT APPLICATION NUMBER: US/10/195,730
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: US/09/281,976
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/060,837
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,862
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 390
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 291
LENGTH: 27
TYPE: PRT
ORGANISM: Homo sapiens
US-10-195-730-291

Query Match 25.6%; Score 33; DB 14; Length 27;
Best Local Similarity 56.2%; Pred. No. 4.2e+02;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 7 INLADRMGLSGVQEI 22
DB 2 IHLEKRSGLSETQII 17

RESULT 11
US-09-798-889-121
Sequence 121, Application US/09798889
Publication No. US20030004324A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 31 Human secreted proteins
FILE REFERENCE: P2026P1
CURRENT APPLICATION NUMBER: US/09/798,889
CURRENT FILING DATE: 2001-03-06
PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-09-09
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-03-12
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-03-12
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-03-12
PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
NUMBER OF SEQ ID NOS: 185

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-889-121

Query Match
Best Local Similarity 25.6%; Score 33; DB 10; Length 28;
Matches 6; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 5 KKNLADRMGLSGVQEIKEQ 25
|||::|||
Db 3 KKIDIPSVFIGESSANSLKDE 23

RESULT 12
US-10-633-680-121
; Sequence 121, Application US/10633680
; Publication No. US2004030115A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 31 Human secreted proteins
; FILE REFERENCE: P2026PIC2
; CURRENT APPLICATION NUMBER: US/10/633,680
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: 09/787,889
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 09/393,022
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: PCT/US99/05721
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/077,714
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/077,686
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/077,687
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/077,696
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 121
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-633-680-121

Query Match
Best Local Similarity 25.6%; Score 33; DB 12; Length 28;
Matches 6; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 5 KKNLADRMGLSGVQEIKEQ 25
|||::|||
Db 3 KKIDIPSVFIGESSANSLKDE 23

RESULT 13
US-09-864-761-42338
; Sequence 42338, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42338
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL138831.2
; OTHER INFORMATION: EXPRESSED IN HEILA, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
; OTHER INFORMATION: EST_HUMAN HIT: AUI20988.1, EVALUE 2.00e-09
; OTHER INFORMATION: SWISSPROT HIT: Q61136, EVALUE 1.00e-10
US-09-864-761-42338

Query Match
Best Local Similarity 25.6%; Score 33; DB 9; Length 29;
Matches 10; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 2 INPKKINLADRMGLSGVQEIKE 24
|||::|||
Db 9 INPTKLLAD---LIGCQLPE 27

RESULT 14
US-10-695-499-147
; Sequence 147, Application US/10695499
; Publication No. US20040126391A1
; GENERAL INFORMATION:
; APPLICANT: Scalato, Enzo
; APPLICANT: Masignani, Vega
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Grandi, Guido
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:40:36 ; Search time 10.6667 Seconds
(without alignments)
225.449 Million cell updates/sec

Title: US-09-171-432A-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 6282

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	25.6	19	2	I52721
2	31	24.0	28	2	S55729
3	30.5	23.6	23	2	I64839
4	30.5	23.6	24	2	S02073
5	29	22.5	17	2	I51910
6	29	22.5	25	2	PC1314
7	28	21.7	19	2	JC2059
8	28	21.7	19	2	PH1315
9	28	21.7	15	2	SI7561
10	27	20.9	12	2	S65626
11	26.5	20.5	27	2	C44636
12	26	20.2	19	2	A61243
13	26	20.2	20	2	B61080
14	26	20.2	20	2	S35970
15	26	20.2	21	2	FX0078
16	26	20.2	21	2	S35978
17	26	20.2	23	2	S35975
18	26	20.2	23	2	S35976
19	26	20.2	30	2	E85694
20	25.5	19.8	26	2	T15472
21	25	19.4	20	2	C49753
22	25	19.4	25	2	D47689
23	25	19.4	25	2	S68005
24	25	19.4	27	2	E44621
25	25	19.4	30	2	S68312
26	25	19.4	30	2	B70165
27	25	19.4	30	2	A36733
28	25	19.4	30	2	H84074
29	24.5	19.0	27	2	B44636

30 24.5 19.0 30 2 S42364
31 24 18.6 19 4 I54264
32 24 18.6 20 2 JP0055
33 24 18.6 22 2 S07394
34 24 18.6 23 2 S70780
35 24 18.6 23 2 I54773
36 24 18.6 23 2 S43469
37 24 18.6 23 2 S43469
38 24 18.6 27 2 A59480
39 24 18.6 28 2 S47624
40 24 18.6 28 2 A41476
41 24 18.6 28 2 A44877
42 24 18.6 30 2 S21195
43 24 18.6 30 2 C95030
44 23.5 18.2 23 2 S45030
45 23 17.8 14 2 G61308

ALIGNMENTS

RESULT 1

I52721

Gene hMLH1 protein - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 21-Jul-2000

C:Accession: I52721

R:Charbonnier, F.; Martin, C.; Scotte, M.; Sibert, L.; Moreau, V.; Frebourg, T.
Cancer Res. 55, 1839-1841, 1995

A:Title: Alternative splicing of MLH1 messenger RNA in human normal cells.

A:Reference number: I52721; MUID:95246019; PMID:7728749

A:Accession: I52721

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-19 <RES>

A:Cross-references: GB:S77856; NID:g998493; PIDN:AAB34135.1; PID:g998494

C:Genetics:

A:Gene: hMLH1

Query Match 25.6%; Score 33; DB 2; Length 19;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKI 7
Db 13 WLNPPQV 19

RESULT 2

S55729

orotidine-5'-monophosphate decarboxylase - Aspergillus awamori (fragments)

C:Species: Aspergillus awamori

C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999

C:Accession: S55729

R:Gouka, R.J.; Heasing, J.G.M.; Stam, H.; Musters, W.; van den Hondel, C.A.M.J.J.
Curr. Genet. 27, 536-540, 1995

A:Title: A novel strategy for the isolation of defined pyrG mutants and the development

A:Reference number: S55729; MUID:96031709; PMID:7553938

A:Accession: S55729

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-18;19-28 <GOU>

C:Superfamily: orotidine-5'-phosphate decarboxylase; orotidine-5'-phosphate decarboxylas

Query Match 24.0%; Score 31; DB 2; Length 28;

Best Local Similarity 55.6%; Pred. No. 5.4e+02;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 7 INLADRLMG 15

Db 14 LDLADRVWG 22

```
RESULT 3
I64839
SP-A1 (alpha, beta, epsilon) and SP-A2 (alpha, beta) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 20-Aug-1999
C:Accession: I64839; I64842; I64841; I64845; I64838
R:McCormick, S.M.; Boggarani, V.; Mendelson, C.R.
Am J. Physiol. 266, L354-L366, 1994
A:Title: Characterization of mRNA transcripts and organization of human SP-A1 and SP-A2
A:Reference number: I51910; MUID:94234365; PMID:8179012
A:Accession: I64839
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-23 <RES>
A:Cross-references: GB:S69681; NID:G546673; PIDN:AAB30731.1; PID:G546674
A:Accession: I64842
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-23 <RES>
A:Cross-references: GB:S69686; NID:G546681; PIDN:AAB30735.1; PID:G546682
A:Accession: I64841
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-23 <RES>
A:Cross-references: GB:S69685; NID:G546679; PIDN:AAB30734.1; PID:G546680
A:Accession: I64845
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-23 <RES>
A:Cross-references: GB:S69690; NID:G546687; PIDN:AAB30738.1; PID:G546688
A:Accession: I64838
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-23 <RES>
A:Cross-references: GB:S69680; NID:G546671; PIDN:AAB30730.1; PID:G546672
C:Genetics:
A:Gene: SP-A2; SP-A1
C:Superfamily: mannose-binding lectin; C-type lectin homology

Query Match 23.6%; Score 30.5; DB 2; Length 23;
Best Local Similarity 37.5%; Pred. No. 5.2e+02;
Matches 9; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

QY 1 WLNPKKINLADRMGLSGVQ-EIK 23
||| :|| :|| :||
Db 2 WLCPLALNLI--LMAAGACEVK 23

RESULT 4
S02073
photosystem II 10K protein precursor - wheat (fragment)
C:Species: Triticum aestivum (common wheat)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 23-Jul-1999
C:Accession: S02073
R:Webber, A.N.; Packman, L.C.; Gray, J.C.
FEBS Lett. 242, 435-438, 1989
A:Title: A 10 kDa polypeptide associated with the oxygen-evolving complex of photosystem II
A:Reference number: S02073; MUID:89121115; PMID:2644135
A:Accession: S02073
A:Molecule type: protein
A:Residues: 1-24 <WEB>
C:Superfamily: photosystem II 10K protein
C:Keywords: chloroplast; photosynthesis; photosystem II; thylakoid

Query Match 23.6%; Score 30.5; DB 2; Length 24;
Best Local Similarity 57.1%; Pred. No. 5.5e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 5 KKNLADRMGLSLG 18
||| :|| :||
Db 4 KKIKV-DKPLGLGL 16
```

```
RESULT 5
I51910
SP-A2 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C:Accession: I51910
R:McCormick, S.M.; Boggarani, V.; Mendelson, C.R.
Am J. Physiol. 266, L354-L366, 1994
A:Title: Characterization of mRNA transcripts and organization of human SP-A1 and SP-A2
A:Reference number: I51910; MUID:94234365; PMID:8179012
A:Accession: I51910
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-17 <RES>
A:Cross-references: GB:S69679; NID:G546669; PIDN:AAB30729.1; PID:G546670
C:Genetics:
A:Gene: SP-A2
C:Superfamily: mannose-binding lectin; C-type lectin homology

Query Match 22.5%; Score 29; DB 2; Length 17;
Best Local Similarity 38.9%; Pred. No. 6.4e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 2; Gaps 1;

QY 1 WLNPKKINLADRMGLSLG 18
||| :|| :||
Db 2 WLCPLALNLI--LMAAG 17

RESULT 6
PC1314
large granule L1 chain - horseshoe crab (Tachyplesus tridentatus) (fragment)
C:Species: Tachyplesus tridentatus
C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
C:Accession: PC1314
R:Shigenaga, T.; Takayenoki, Y.; Kawasaki, S.; Seki, N.; Toh, Y.; Ito, A.; Iwa
J. Biochem. 114, 307-316, 1993
A:Title: Separation of large and small granules from horseshoe crab (Tachyplesus tridentatus)
A:Reference number: PC1309; MUID:94110249; PMID:8282718
A:Accession: PC1314
A:Molecule type: protein
A:Residues: 1-25 <SHI>
C:Comment: This protein participates in immobilization of invading microbes.

Query Match 22.5%; Score 29; DB 2; Length 25;
Best Local Similarity 46.7%; Pred. No. 9.6e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 NLADRMGLSLGVQEI 22
||| :|| :||
Db 5 NKMEEMLXDSGDQEV 19

RESULT 7
JC2059
homeobox 4 protein - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 23-Mar-1995
C:Accession: JC2059
R:Feng, X.H.; Kung, S.D.
Biochem. Biophys. Res. Commun. 198, 1012-1019, 1994
A:Title: Identification of differentially expressed members of tobacco homeobox families
A:Reference number: JC2057; MUID:94161708; PMID:7509595
A:Accession: JC2059
A:Molecule type: DNA
A:Residues: 1-19 <FEN>
A:Experimental source: leaf
C:Genetics:
A:Gene: Hot4
C:Keywords: homeobox

Query Match 21.7%; Score 28; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```


Job time : 10.6667 secs

A;Title: 5-carboxymethyl-2-hydroxy-muconic semialdehyde dehydrogenases of Escherichia coli
A;Reference number: A61080
A;Accession: B61080
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <FAW>
C;Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
C;Keywords: oxidoreductase

Query Match 20.2%; Score 26; DB 2; Length 20;
Best Local Similarity 36.4%; Pred. No. 2.1e+03;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WLNPKKINLAD 11
||| : :
Db 7 WINGKNVAGSD 17

RESULT 14

S35970
ribosomal protein L10 - Citrobacter freundii (fragment)
C;Species: Citrobacter freundii
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Oct-1997
C;Accession: S35970

R;Thyvoloup, A.
submitted to the EMBL Data Library, August 1993

A;Reference number: S35970
A;Accession: S35970
A;Molecule type: DNA
A;Residues: 1-20 <ZHY>
A;Cross-references: EMBL:X74448

C;Genetics:

A;Gene: rplJ

C;Superfamily: Escherichia coli ribosomal protein L10

C;Keywords: protein biosynthesis; ribosome

Query Match 20.2%; Score 26; DB 2; Length 20;
Best Local Similarity 31.2%; Pred. No. 2.1e+03;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 7 INLADRMGLSGVQEI 22
:||| : : :
Db 3 INLQDKQAIIVAEVSEV 18

RESULT 15

PX0078
alanine dehydrogenase (EC 1.4.1.1) [validated] - Phormidium lapideum (fragment)

C;Species: Phormidium lapideum

C;Date: 22-Apr-1995 #sequence_revision 26-May-1995 #text_change 20-Oct-2000

C;Accession: PX0078

R;Sawa, Y.; Tani, M.; Murata, K.; Shibata, H.; Ochiai, H.

J. Biochem. 116, 995-1000, 1994

A;Title: Purification and characterization of alanine dehydrogenase from a cyanobacterium

A;Reference number: PX0078; MUID:95204408; PMID:7896761

A;Accession: PX0078

A;Molecule type: protein

A;Residues: 1-21 <SAW>

C;Function:

A;Description: catalyzes the oxidative deamination of L-alanine to pyruvate and ammonia

C;Superfamily: alanine dehydrogenase; alanine dehydrogenase homology

C;Keywords: oxidoreductase

Query Match 20.2%; Score 26; DB 2; Length 21;
Best Local Similarity 42.9%; Pred. NO. 2.3e+03;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 PKKINLADRMGLSL 17
||| : : :
Db 6 PREINKQEFRRVGLS 19

Search completed: July 15, 2004, 14:52:15

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 14:37:01 ; Search time 6.66667 Seconds
(without alignments)
195.263 Million cell updates/sec

Title: US-09-171-432A-48

Perfect score: 129

Sequence: 1 MLNPKKINLADRMGLSGVQETKEQ 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 2073

Minimum DB seq length: 0

Maximum DB seq length: 30

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	30.5	23.6	24	1	PSBR_WHEAT	P12358	triticum ae
2	28	21.7	25	1	UBLI_BOVIN	P23356	bos taurus
3	26	20.2	19	1	RL10_CITR	P43448	citrobacter
4	26	20.2	20	1	RL10_PROVU	P51411	proteus vul
5	26	20.2	22	1	RL10_KLEPN	P41190	klebsiella
6	26	20.2	23	1	RL10_ENTCL	Q47608	enterobacte
7	25	19.4	13	1	IDHC_FIG	P20304	sus scrofa
8	25	19.4	19	1	LPGE_ECOLI	P33236	escherichia
9	25	19.4	20	1	YPRB_SERMA	P22581	seriatia ma
10	25	19.4	25	1	FLB3_TREHY	P80161	treponema h
11	25	19.4	27	1	MDO2_PSEMR	P83447	pseudonana
12	25	19.4	30	1	Y523_BORBU	O51473	borrelia bu
13	25	19.4	30	1	Y523_BORBU	P24244	escherichia
14	24	18.6	22	1	DNIV_EPD10	Q38199	bacterioph
15	24	18.6	23	1	NDF2_ARTSK	P83576	arthrobacte
16	24	18.6	28	1	ETX2_BACCE	P80568	bacillus ce
17	24	18.6	29	1	ATP9_PICPT	Q06838	pichia pljp
18	24	18.6	29	1	GALA_ONCMY	P47213	oncorhynch
19	23.5	18.2	19	1	FLA3_SPIAU	P21986	spirochaeta
20	23.5	18.2	27	1	FLA4_SPIAU	P21987	spirochaeta
21	23	17.8	12	1	CD11_LITXA	P56245	ash yellow
22	23	17.8	13	1	RS19_ASHYP	O44592	ash yellow
23	23	17.8	15	1	UBLI_MONDO	P50103	monodelphis
24	23	17.8	19	1	FM01_CAVPO	P49328	cavia porce
25	23	17.8	19	1	UP27_UPEMJ	P82039	uperoleia m
26	23	17.8	25	1	CR22_LITGI	P56234	litoria gil
27	23	17.8	25	1	CR25_LITGI	P56237	litoria gil
28	23	17.8	27	1	MDH_RHORI	P80459	rhodospiril
29	23	17.8	30	1	CXEX_CONCN	P58928	conus conso
30	23	17.8	30	1	PGG3_PAGCO	P82416	pachycondyl
31	22.5	17.4	23	1	PQQA_ENTIT	P59726	enterobacte
32	22.5	17.4	23	1	PQQA_KLEPN	P27503	klebsiella
33	22	17.1	10	1	TKS1_AEDAE	P42634	aedes aegypt

RESULT 1

PSBR_WHEAT

ID	PSBR_WHEAT	STANDARD	PRT	24 AA
AC	P12358			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Photosystem II 10 kDa polypeptide (Fragment).			
GN	PSBR.			
OS	Triticum aestivum (Wheat).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;			
OC	Triticeae; Triticum.			
OX	NCBI_TaxID=4565;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=89121115; PubMed=2644135;			
RA	Webber A.N., Packman L.C., Gray J.C.;			
RT	"A 10 kDa polypeptide associated with the oxygen-evolving complex of			
RT	photosystem II has a putative C-terminal non-cleavable thylakoid			
RT	transfer domain."			
RL	FEBS Lett. 242:435-438 (1989).			
CC	!- FUNCTION: Associated with the oxygen-evolving complex of			
CC	photosystem II.			
CC	!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated			
CC	with the photosystem II complex.			
CC	!- SIMILARITY: Belongs to the psbr family.			
DR	PIR; S02073; S02073.			
DR	InterPro; IPR006814; Psbr.			
DR	Pfam; PF04725; Psbr; 1.			
KW	Photosynthesis; Photosystem II; Chloroplast; Thylakoid; Membrane.			
FT	NON_TER 24 24			
SQ	SEQUENCE 24 AA; 2339 MW; 4D7F56C797E1503E CRC64;			

Query Match 23.6%; Score 30.5; DB 1; Length 24;
Best Local Similarity 57.1%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 5 KKINLADRMGLSLG 18
||| : |||
Db 4 KKIKV-DKPLGLG 16

RESULT 2

UBLI_BOVIN

ID	UBLI_BOVIN	STANDARD	PRT	25 AA
AC	P23356			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH-			
DE	L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5)			
DE	(PGP 9.5) (PGP9.5) (Fragment).			
GN	UCHL1.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein.
FT INIT MET 0 BY SIMILARITY.
RN NCBI_TaxID=9913;
RN SEQUENCE 19 AA; 2026 MW; E20A02CD7851DB16 CRC64;
RC TISSUE=Brain;
RX MEDLINE=92008646; PubMed=1833240;
RA Giambanco I., Bianchi R., Ceccarelli P., Pula G., Sorci G.,
RA Antonoli S., Bocchini V., Donato R.;
RT 'Neuron-specific' protein gene product 9.5 (PGP 9.5) is also
RT expressed in glioma cell lines and its expression depends on cellular
RT growth state.;
RL FEBS Lett. 290:131-134 (1991).
CC -!- FUNCTION: Ubiquitin-protein hydrolase is involved both in the
CC processing of ubiquitin precursors and of ubiquitinated proteins.
CC This enzyme is a thiol protease that recognizes and hydrolyzes
CC a peptide bond at the C-terminal glycine of ubiquitin.
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Neurons and cells of the diffuse
CC neuroendocrine system and their tumors.
CC -!- SIMILARITY: Belongs to peptidase family C12.
DR PIR; S17561; S17561.
DR MEROPS; C12.001; -.
DR InterPro; IPR001578; Peptidase C12.
DR Pfam; PF01088; Peptidase C12; 1.
DR PROSITE; PS00140; UCH 1; PARTIAL.
KW Ub1 conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT NON TER 25
FT SEQUENCE 25 AA; 2812 MW; 26BB5ADD0A754D55 CRC64;
Query Match 21.7%; Score 28; DB 1; Length 25;
Best Local Similarity 29.4%; Pred. No. 6.6e+02;
Matches 5; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
QY 2 LNPKKINLADRMGLSG 18
DB 8 IDPEMLNKVLRGVAG 24
RESULT 3
RL10 CITFR STANDARD; PRT; 19 AA.
ID AC P43448;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10 (Fragment).
GN RPLJ.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN SEQUENCE FROM N.A.
RP STRAIN=771;
RC Zhyvoloup A.N.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74448; CAA52459.1; -.
DR PIR; S35970; S35970.
DR HAMAP; MF 00362; -; 1.
DR InterPro; IPR002363; Ribosomal_L10aeb.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein.
FT INIT MET 0 BY SIMILARITY.
FT NON TER 20
FT SEQUENCE 20 AA; 2083 MW; 76520A02CD7851DB CRC64;
Query Match 20.2%; Score 26; DB 1; Length 20;
Best Local Similarity 31.2%; Pred. No. 1.1e+03;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 7 INLADRMGLSGVQEI 22
DB 2 LNLQDKQAIVAEVSEV 17
RESULT 4
RL10 PROUV STANDARD; PRT; 20 AA.
ID AC P51411;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10 (Fragment).
GN RPLJ.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Proteus.
OX NCBI_TaxID=585;
RN SEQUENCE FROM N.A.
RP Zhyvoloup A.N.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74446; CAA52457.1; -.
DR PIR; S35978; S35978.
DR HAMAP; MF 00362; -; 1.
DR InterPro; IPR002363; Ribosomal_L10eub.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein.
FT INIT MET 0 BY SIMILARITY.
FT NON TER 20
FT SEQUENCE 20 AA; 2083 MW; 76520A02CD7851DB CRC64;
Query Match 20.2%; Score 26; DB 1; Length 20;
Best Local Similarity 31.2%; Pred. No. 1.1e+03;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 7 INLADRMGLSGVQEI 22
DB 2 LNLQDKQAIVAEVSEV 17
RESULT 5
RL10 KLEPN STANDARD; PRT; 22 AA.
ID AC P41190;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10 (fragment).
GN RPLJ.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92008646; PubMed=1833240;
RA Giambanco I., Bianchi R., Ceccarelli P., Pula G., Sorci G.,
RA Antonoli S., Bocchini V., Donato R.;
RT 'Neuron-specific' protein gene product 9.5 (PGP 9.5) is also
RT expressed in glioma cell lines and its expression depends on cellular
RT growth state.;
RL FEBS Lett. 290:131-134 (1991).
CC -!- FUNCTION: Ubiquitin-protein hydrolase is involved both in the
CC processing of ubiquitin precursors and of ubiquitinated proteins.
CC This enzyme is a thiol protease that recognizes and hydrolyzes
CC a peptide bond at the C-terminal glycine of ubiquitin.
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Neurons and cells of the diffuse
CC neuroendocrine system and their tumors.
CC -!- SIMILARITY: Belongs to peptidase family C12.
DR PIR; S17561; S17561.
DR MEROPS; C12.001; -.
DR InterPro; IPR001578; Peptidase C12.
DR Pfam; PF01088; Peptidase C12; 1.
DR PROSITE; PS00140; UCH 1; PARTIAL.
KW Ub1 conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT NON TER 25
FT SEQUENCE 25 AA; 2812 MW; 26BB5ADD0A754D55 CRC64;
Query Match 21.7%; Score 28; DB 1; Length 25;
Best Local Similarity 29.4%; Pred. No. 6.6e+02;
Matches 5; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
QY 2 LNPKKINLADRMGLSG 18
DB 8 IDPEMLNKVLRGVAG 24
RESULT 3
RL10 CITFR STANDARD; PRT; 19 AA.
ID AC P43448;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10 (Fragment).
GN RPLJ.
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN SEQUENCE FROM N.A.
RP STRAIN=771;
RC Zhyvoloup A.N.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74448; CAA52459.1; -.
DR PIR; S35970; S35970.
DR HAMAP; MF 00362; -; 1.
DR InterPro; IPR002363; Ribosomal_L10aeb.

```

OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 5054;
RA Zhyvoloup A.N.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74445; CAA52455.1; -
DR PIR; S35975; S35975.
DR HAMAP; MF_00362; -; 1.
DR InterPro; IPR002363; Ribosomal_L10eub.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein.
FT INIT_MET 0 0 BY SIMILARITY.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2268 MW; 113D96520A02CD78 CRC64;

Query Match
Best Local Similarity 20.2%; Score 26; DB 1; Length 22;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 7 INLADRLMLGLSGVQEI 22
Db :||| : : : :
2 LNLQDKQAIVAEVSEV 17

RESULT 6
ID RL10 ENTCL STANDARD; PRT; 23 AA.
AC Q47608;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L10 (Fragment).
GN RPLJ.
OS Enterobacter cloacae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
OX NCBI_TaxID=550;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhyvoloup A.;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74444; CAA52455.1; -
DR HAMAP; MF_00362; -; 1.
DR InterPro; IPR002363; Ribosomal_L10eub.
DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein.
FT INIT_MET 0 0 BY SIMILARITY.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2355 MW; 39413D96520A02CD CRC64;

Query Match
Best Local Similarity 20.2%; Score 26; DB 1; Length 23;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 7 INLADRLMLGLSGVQEI 22
Db :||| : : : :
2 LNLQDKQAIVAEVSEV 17

Best Local Similarity 31.2%; Pred. No. 1.2e+03;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 7 INLADRLMLGLSGVQEI 22
Db :||| : : : :
2 LNLQDKQAIVAEVSEV 17

RESULT 7
ID IDHC_PIG STANDARD; PRT; 13 AA.
AC P20304;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Isocitrate dehydrogenase [NADP] cytoplasmic (EC 1.1.1.42)
DE (Oxalosuccinate decarboxylase) (IDH) (NADP+-specific ICDH) (IDP)
DE (Fragment).
DE IDH1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE
RX MEDLINE=87308293; PubMed=2887570;
RA Bailey J.M., Colman R.F.;
RT "Isolation of the glutamyl peptide labeled by the nucleotide analogue
RT 2-(4-bromo-2,3-dioxobutylthio)-1,N(6)-ethenoadenosine
RT 2',5'-biphosphate in the active site of NADP+-specific isocitrate
RT dehydrogenase."
RL J. Biol. Chem. 262:12614-12619(1987).
CC -!- CATALYTIC ACTIVITY: Isocitrate + NADP(+) = 2-oxoglutarate + CO(2)
CC + NADPH.
CC -!- SUBUNIT: Homodimer.
DR InterPro; IPR001804; Isodh.
DR PROSITE; PS00470; IDH_IMDH; PARTIAL.
KW Oxidoreductase; NADP; Tricarboxylic acid cycle.
FT NON_TER 1 1
FT ACT_SITE 5 5 POTENTIAL.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1353 MW; 1B640F0E9F7C71E0 CRC64;

Query Match
Best Local Similarity 19.4%; Score 25; DB 1; Length 13;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 NLADRLMLGLSGVQ 20
Db :||| : : : :
1 DLAGEIHGLSNVK 13

RESULT 8
ID LPGE_ECOLI STANDARD; PRT; 19 AA.
AC P33236;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gef leader peptide.
DE GEFL OR B0018.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;

```

RN SEQUENCE FROM N.A.
 RP MEDLINE=92048481; PubMed=1943701;
 RX Poulson L.K., Refn A., Molin S., Andersson P.;
 RA "The *gef* gene from *Escherichia coli* is regulated at the level of
 RT translation.";
 RL Mol. Microbiol. 5:1639-1648 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474 (1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000112; AAC73129.1; ALT_TERM.
 DR EcoGene; EG12074; *gefL*.
 KW Leader peptide; Complete proteome.
 SQ SEQUENCE 19 AA; 2259 MW; 19B3EDF371EB0BEB CRC64;
 CC -----
 Query Match 19.4%; Score 25; DB 1; Length 19;
 Best Local Similarity 25.2%; Pred. No. 1.4e+03;
 Matches 7; Conservative 5; Mismatches 4; Indels 8; Gaps 1;
 QY 2 LNPKKINLADRMGLSGVQEIKEQ 25
 DB 2 LNTCRVPLTDR-----KVKEK 17
 CC -----
 RESULT 9
 YPRB SERMA
 ID YPRB SERMA STANDARD; PRT; 20 AA.
 AC P22581;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein in *proB* 5' region (Fragment).
 OS *Serratia marcescens*.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; *Serratia*.
 CC NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sr41;
 RX MEDLINE=91237315; PubMed=1851803;
 RA Omori K., Suzuki S., Imai Y., Komatsubara S.;
 RA "Analysis of the *Serratia marcescens* *proB*A operon and feedback
 RT control of proline biosynthesis.";
 RL J. Gen. Microbiol. 137:509-517 (1991).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D90351; BAAL4363.1; -;
 DR EMBL; X53086; CAA37253.1; -;
 DR PIR; C49753; C49753.
 KW Hypothetical protein.

FT NON TER 1
 SQ SEQUENCE 20 AA; 2248 MW; 4DD7777735276674 CRC64;
 Query Match 19.4%; Score 25; DB 1; Length 20;
 Best Local Similarity 41.7%; Pred. No. 1.5e+03;
 Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 14 LGLSGVQEIKEQ 25
 DB 3 LGLEPADDFKEK 14
 CC -----
 RESULT 10
 FLB3 TREHY
 ID FLB3 TREHY STANDARD; PRT; 25 AA.
 AC P80161;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Flagellar filament core protein *flaB3* (32 kDa core protein)
 DE (Fragment).
 DE DE
 GN FLAB3.
 OS *Treponema hyodysenteriae* (Serpulina hyodysenteriae).
 CC Bacteria; Spirochaetes; Spirochaetales; Brachyspiraceae; Brachyspira.
 CC NCBI_TaxID=159;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=C5;
 RX MEDLINE=93139764; PubMed=1487733;
 RA Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
 RA van der Zeijst B.A.M., Kusters J.G.;
 RT "The periplasmic flagella of *Serpulina* (*Treponema*) *hyodysenteriae* are
 RT composed of two sheath proteins and three core proteins.";
 RL J. Gen. Microbiol. 138:2697-2706 (1992).
 CC -!- FUNCTION: Component of the core of the flagella.
 CC -!- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF TWO
 CC SHEATH PROTEINS, FLAA1 (44 kDa) AND FLAA2 (35 kDa) AROUND A CORE
 CC THAT CONTAINS THREE PROTEINS FLAB1 (37 kDa), FLAB2 (34 kDa) AND
 CC FLAB3 (32 kDa).
 CC -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
 CC -!- SIMILARITY: Belongs to the bacterial flagellin family.
 DR PIR; D47689; D47689.
 KW Flagellum; Periplasmic.
 FT NON TER 25
 SQ SEQUENCE 25 AA; 2746 MW; 1AAA0CA722EE3D0A CRC64;
 Query Match 19.4%; Score 25; DB 1; Length 25;
 Best Local Similarity 55.6%; Pred. No. 1.9e+03;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 10 ADRMLGLSG 18
 DB 12 ANRQLNLTG 20
 CC -----
 RESULT 11
 MD02 PSEMR
 ID MD02 PSEMR STANDARD; PRT; 27 AA.
 AC P83477;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Macrodontain II (EC 3.4.22.-) (Fragment).
 OS *Pseudananas macrodantes* (*Pseudananas sagenarius*).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Bromeliaceae;
 CC Pseudananas.
 CC NCBI_TaxID=203992;
 RN [1]
 RP SEQUENCE, CATALYTIC ACTIVITY, ENZYME REGULATION, SUBUNIT, AND MASS
 RP SPECTROMETRY.
 RC TISSUE=Fruit.
 RX MEDLINE=21408736; PubMed=11517945;

```

Query Match      19.4%; Score 25; DB 1; Length 30;
Best Local Similarity 25.0%; Pred. No. 2.3e+03;
Matches 3; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY      14 LGLSGVQEIKEQ 25
          :|::|::|::|
DB       1 MGFGIIALKQR 12

RESULT 13
YCCB_ECOLI
ID_YCCB_ECOLI   STANDARD;           PRT;    30 AA.
AC      P2424;
DT      01-MAR-1992 (Rel. 21, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical protein yccb.
YCBB OR B0979.1.
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
NCBI_TaxID=562;
[1]
RN      SEQUENCE FROM N.A.
RP      STRAIN=K12.
RC      MEDLINE=90368616; PubMed=2168385;
RX      Dassa J., Marck C., Boquet P.L.;
RA      "The complete nucleotide sequence of the Escherichia coli gene appA
RT      reveals significant homology between pH 2.5 acid phosphatase and
RL      glucose-1-phosphatase.";
RJ      J. Bacteriol. 172:5497-5500(1990).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12;
RX      MEDLINE=92049231; PubMed=1658595;
RA      Dassa J., Fsihi H., Marck C., Dion M., Kieffer-Bontemps M.,
RB      Boquet P.L.;
RT      "A new oxygen-regulated operon in Escherichia coli comprises the
RT      genes for a putative third cytochrome oxidase and for pH 2.5 acid
RT      phosphatase (appA)".;
RL      Mol. Gen. Genet. 229:341-352(1991).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12." ;
RT      Science 277:1453-1474(1997).
RC      -!- SIMILARITY: TO E.COLI YBGT.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M58708; AAA72085.1; -.
CDR     DR
CDR     EMBL; S63811; AAB20287.1; -.
CDR     EMBL; AE000199; -. NOT_ANNOTATED_CDS.
CDR     PIR; A36733; A36733.
CDR     EcGene; EG11114; yccb.
KW      Hypothetical protein; Transmembrane; Inner membrane;
KW      Complete proteome.
TRANSMEM      4      24      POTENTIAL.
SEQUENCE      30 AA; 3597 MW; 9E27FCD533329606 CRC64;

Query Match      19.4%; Score 25; DB 1; Length 30;

```

```
Best Local Similarity 60.0%; Pred. No. 2.3e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPK 5
DB 23 WLDPR 27

RESULT 14
DNV BPD10 STANDARD; PRT; 22 AA.
AC Q38139;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-invertase (Fragment).
GN GIN.
OS Bacteriophage D108.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC Mu-like viruses.
OC NCHI_TaxID=10671;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=87316928; PubMed=2957646;
RX Szatmari G.B., Lapointe M., Dubow M.S.;
RA "The right end of transposable bacteriophage D108 contains a 520 base
RT pair protein-encoding sequence not present in bacteriophage Mu.";
RL Nucleic Acids Res. 15:6691-6703(1987).
CC -1- FUNCTION: THIS PROTEIN CATALYZES THE INVERSION OF A 3000-BP
CC SEGMENT OF PHAGE DNA. THE INVERSION RESULTS IN A MODIFICATION OF
CC THE 3' END OF THE TAIL FIBER GENE AND ALTERS THE HOST SPECIFICITY.
CC -1- SIMILARITY: Belongs to the site-specific recombinase resolvase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL: X05926; CAA29365.1; -.
PIR: S07394; S07394.
InterPro: IPR006118; Recombinase.
DR PROSITE; PS00397; RECOMBINASES 1; PARTIAL.
DR PROSITE; PS00398; RECOMBINASES 2; PARTIAL.
KW DNA recombination; DNA integration; DNA-binding; DNA invertase.
FT NON TER 1
FT DNA_BIND <1 5 H-T-H MOTIF (PROBABLE).
SQ SEQUENCE 22 AA; 2748 MW; 097B607032767C38 CRC64;

Query Match 18.6%; Score 24; DB 1; Length 22;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 NPKINLADR 12
DB 13 NDDRINQIDR 22

RESULT 15
NDP2 ARTSK STANDARD; PRT; 23 AA.
AC P83576;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NADP phosphatase II (EC 3.1.3.-) (Fragment).
OS Arthrobacter sp. (strain KM).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=184230;
```

```
[1]
RN SEQUENCE, SUBUNIT, AND SUBCELLULAR LOCATION.
RP Kawai S., Mori S., Mukai T., Murata K.;
RA "Properties of NADP phosphatase I and II of Arthrobacter sp. strain
RT KM.";
RL Eur. J. Biochem. 0:0-0(2003).
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: Maximal activity at pH 7-8.
DR InterPro: IPR000437; Prok_lipoprot_S.
KW Hydrolase.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2453 MW; 19547F56E0655886 CRC64;

Query Match 18.6%; Score 24; DB 1; Length 23;
Best Local Similarity 40.0%; Pred. No. 2.5e+03;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 KXINLADRL 14
DB 6 QKANLGNQMI 15

Search completed: July 15, 2004, 14:47:36
Job time : 7.66667 secs
```


Result No.	Query			Description	
	Score	Match	Length	ID	
1	34	26.4	27	5	Q9BM49 moniliformi
2	33	25.6	19	4	Q16335 homo sapien
3	32	24.8	25	5	Q25183 hydra atten
4	32	24.8	27	5	P91713 dugesia tig
5	31	24.0	26	17	Q8Z2V3 pyrobaculum
6	31	24.0	27	5	P91901 beroe ovata
7	31	24.0	27	5	Q9BM48 moniliformi
8	31	24.0	27	13	Q90297 carassius a
9	31	24.0	30	8	Q9T2P9 narcissus p
10	30.5	23.6	27	4	P78489 homo sapien
11	30.5	23.6	27	5	Q61676 drosophila
12	30	23.3	14	13	Q73591 gallus gall
13	30	23.3	17	2	Q9X313 prochloroco
14	30	23.3	21	2	Q9X3L1 prochloroco
15	30	23.3	21	2	Q9X3K2 prochloroco
16	30	23.3	24	5	Q9U9B7 ceratitlis c

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HMLH1 protein (Fragment).
GN HMLH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95246019; PubMed=7728749;
RA Charbonnier F., Martin C., Scotte M., Sibert L., Moreau V.,
RA Frebourg T.;
RT "Alternative splicing of MLH1 messenger RNA in human normal cells.";
RL Cancer Res. 55:1839-1841(1995).
DR EMBL; S77856; AAB34135.1; -;
DR PIR; I52721; I52721.
FT NON_TER 1
SQ SEQUENCE 19 AA; 2228 MW; 6B7CC17D7F651B74 CRC64;
Query Match 25.6%; Score 33; DB 4; Length 19;
Best Local Similarity 57.1%; Pred. No. 5.5e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 WLNPKKI 7
Db 13 WLNPOQV 19
RESULT 3
Q25183 ID Q25183 PRELIMINARY; PRT; 25 AA.
AC Q25183;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CHV2 protein (Fragment).
GN CNH2.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93321609; PubMed=8101168;
RA Shenk M.A., Bode H.R., Steele R.E.;
RT "Expression of Cnox-2, a HOM/HOX homeobox gene in hydra, is correlated
RL with axial pattern formation.";
RL Development 117:657-667(1993).
DR EMBL; M62871; AAA29208.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3144 MW; 90936F858F24D93C CRC64;
Query Match 24.8%; Score 32; DB 5; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy 5 KKINLADRL 14
Db 15 KKLNLSEQI 24
RESULT 4
P91713 ID P91713 PRELIMINARY; PRT; 27 AA.
AC P91713;

DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeodomain protein (Fragment).
GN DTHOXB.
OS Dugesia tigrina (Planarian).
OC Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC Paludicola; Dugesidae; Girardia.
OX NCBI_TaxID=6162;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97158715; PubMed=9006075;
RA Bayascas J.R., Castillo E., Munoz-Marmol A.M., Salo E.B.;
RT "Planarian Hox genes: novel patterns of expression during
RT regeneration.";
RL Development 124:141-148(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99016215; PubMed=9799427;
RA Bayascas J.R., Castillo E., Salo E.B.;
RT "Planhelminthes have a Hox code differentially activated during
RT regeneration, with genes closely related to those of spiral
RT protostomes.";
RL Dev. Genes Evol. 208:467-473(1998).
DR EMBL; X95415; CAA64695.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
FT NON_TER 1
FT NON_TER 27
FT NON_TER 27
SQ SEQUENCE 27 AA; 3339 MW; 9B4F2D5E657EB7F8 CRC64;
Query Match 24.8%; Score 32; DB 5; Length 27;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 4 PKKINLADRM 13
Db 9 PRIDLANRL 18
RESULT 5
Q8ZV3 ID Q8ZV3 PRELIMINARY; PRT; 26 AA.
AC Q8ZV3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE0064a.
GN PAE0064a.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AB009748; AAL62536.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 26 AA; 2611 MW; 9E70055D547FDABF CRC64;
Query Match 24.0%; Score 31; DB 17; Length 26;
Best Local Similarity 35.3%; Pred. No. 1.6e+03;
Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

```

QY      2 LNPCKINLADRMGLSLG 18
Db      4 VGPCKTITFEVIALDG 20

RESULT 6
ID P91901 PRELIMINARY; PRT; 27 AA.
AC P91901;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Homeobox protein BHOx26 (Fragment).
OS Beroe ovata.
OC Eukaryota; Metazoa; Ctenophora; Cyclocoela; Beroidea; Beroe.
OX NCBI_TaxID=10201;
RN [1]
RP SEQUENCE FROM N.A.
RA Witchei H.J., Meech R.W.;
RT "BHOx26 - Beroe Homeobox.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U89381; AAB49471.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3338 MW; 55318651A81607BD CRC64;

Query Match 24.0%; Score 31; DB 5; Length 27;
Best Local Similarity 60.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      5 KKNILADRMRL 14
Db      16 KSNLTDRQV 25

RESULT 7
ID Q9BM48 PRELIMINARY; PRT; 27 AA.
AC Q9BM48;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE LINE-like reverse transcriptase (Fragment).
OS Moniliiformis moniliiformis.
OC Eukaryota; Metazoa; Acanthocephala; Archiacanthocephala;
OC Moniliiformida; Moniliiformidae; Moniliiformis.
OX NCBI_TaxID=10237;
RN [1]
RP SEQUENCE FROM N.A.
RA TRANSPON-LPT-L5 retrotransposon;
RC MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and ancient asexual taxa.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY013946; AAG59930.1; -.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3105 MW; D5D29937B437CDDC CRC64;

Query Match 24.0%; Score 31; DB 5; Length 27;
Best Local Similarity 31.8%; Pred. No. 1.7e+03;
Matches 7; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY      2 LNPCKINLADRMGLSLGVQEI 23

```

```

Db      3 LGPTLFNIYIREVSLSNRQDVQ 24

RESULT 8
ID Q90297 PRELIMINARY; PRT; 27 AA.
AC Q90297;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Homeobox (Fragment).
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=93219357; PubMed=8096640;
RA Levine E.M., Schechter N.;
RT "Homeobox genes are expressed in the retina and brain of adult goldfish.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2729-2733(1993).
DR EMBL; L09693; AAA49178.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
FT NON_TER 1
FT NON_TER 27
SQ SEQUENCE 27 AA; 3338 MW; 55318651A81607BD CRC64;

Query Match 24.0%; Score 31; DB 13; Length 27;
Best Local Similarity 60.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      5 KKNILADRMRL 14
Db      16 KSNLTDRQV 25

RESULT 9
ID Q9T2P9 PRELIMINARY; PRT; 30 AA.
AC Q9T2P9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Heat shock protein 60 (Fragment).
OS Narcissus pseudonarcissus (Daffodil).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae;
OC Narcissus.
OX NCBI_TaxID=39639;
RN [1]
RP SEQUENCE.
RX MEDLINE=96291727; PubMed=8754888;
RA Bonk M., Tadros M., Vandekerckhove J., Al-Babili S., Beyer P.;
RT "Purification and characterization of chaperonin 60 and heat-shock protein 70 from chromatoplasts of Narcissus pseudonarcissus.";
RL Plant Physiol. 111:931-939(1996).
DR HSP; P06139; 1AON.
DR InterPro; IPR008950; GroEL-ATPase.
SQ SEQUENCE 30 AA; 3233 MW; AF5AF69899CE2851 CRC64;

Query Match 24.0%; Score 31; DB 8; Length 30;
Best Local Similarity 36.8%; Pred. No. 1.9e+03;
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 6 KINLADRLGLSGVQEIKE 24
DB 6 KFGVEARALMLRGVEELAD 24

RESULT 10
ID P78489 PRELIMINARY; PRT; 23 AA.
AC P78489;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE SP-A1 epsilon (Fragment).
GN SP-A1 OR SP-A2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94234365; PubMed=8179012;
RA McCormick S.M., Boggarlam V., Mendelson C.R.;
RT "Characterization of mRNA transcripts and organization of human SP-A1
RT and SP-A2 genes.";
RL Am. J. Physiol. 266:L354-L366 (1994).
DR EMBL; S69690; AAB30738.1; -
DR EMBL; S69680; AAB30730.1; -
DR EMBL; S69681; AAB30731.1; -
DR EMBL; S69685; AAB30734.1; -
DR EMBL; S69686; AAB30735.1; -
DR PIR; I64839; I64839.23
FT NON TER 23
SQ SEQUENCE 23 AA; 2419 MW; 8C4FE026553AAB02C CRC64;

Query Match 23.6%; Score 30.5; DB 4; Length 23;
Best Local Similarity 37.5%; Pred. No. 1.7e+03;
Matches 9; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

QY 1 WLNPKKINLADRLGLSGVQ-EIK 23
DB 2 WLCPLAUNLI--LMAASGACEVK 23

RESULT 11
ID Q61676 PRELIMINARY; PRT; 27 AA.
AC Q61676;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE HOX11-D125 protein (D125) (Fragment).
GN HOX11-D125 OR D125.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281593; PubMed=8099440;
RA Dear T.N., Sanchez-Garcia I., Rabbitts T.H.;
RT "The HOX11 gene encodes a DNA-binding nuclear transcription factor
RT belonging to a distinct family of homeobox genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:4431-4435 (1993).
DR EMBL; L08618; AAA28612.1; -
DR FlyBase; FBgn0010394; Hox11-D125.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.

```

```

FT NON TER 1 1
FT NON TER 27 27
SQ SEQUENCE 27 AA; 2937 MW; CABEC2F876F2B3A3 CRC64;

Query Match 23.6%; Score 30.5; DB 5; Length 27;
Best Local Similarity 42.3%; Pred. No. 2e+03;
Matches 11; Conservative 2; Mismatches 8; Indels 5; Gaps 1;

QY 2 LNPCKINLADR-----MLGLSGVQEI 22
DB 1 LYQKYLSPADRDDEIAASLGLSNAQVI 26

RESULT 12
ID O73591 PRELIMINARY; PRT; 14 AA.
AC O73591;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hox C10 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=White leghorn;
RC MEDLINE=98141813; PubMed=9473273;
RA Peale F.V. Jr., Mason K., Hunter A.W., Bothwell M.;
RT "Multiplex display polymerase chain reaction amplifies and resolves
RT related sequences sharing a single moderately conserved domain.";
RL Anal. Biochem. 256:158-168 (1998).
DR EMBL; U34614; AAC36452.1; -
FT NON TER 1 1
FT NON TER 14 14
SQ SEQUENCE 14 AA; 1673 MW; 81258FC9E81FDA45 CRC64;

Query Match 23.3%; Score 30; DB 13; Length 14;
Best Local Similarity 75.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 KKNILADR 12
DB 6 KSNLTDR 13

RESULT 13
ID Q9X3I3 PRELIMINARY; PRT; 17 AA.
AC Q9X3I3;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE Cytochrome b (Fragment).
GN PBTB.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630 (1998).
DR EMBL; AF070176; AAD20791.1; -
FT NON TER 1 1
SQ SEQUENCE 17 AA; 1953 MW; 630843039ADD51B4 CRC64;

Query Match 23.3%; Score 30; DB 2; Length 17;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;

```

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 LADRMGLSG 18
| | | | |
Db 6 LXDRKQGISG 15

RESULT 14

Q9X3L1 PRELIMINARY; PRT; 21 AA.
AC Q9X3L1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE Cytochrome b6/f complex subunit IV (Fragment).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream."
RL Limnol. Oceanog. 43:1615-1630(1998).
DR EMBL; AF070204; AAD23251.1; -.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2306 MW; CC947A26F03C4D67 CRC64;

Query Match 23.3%; Score 30; DB 2; Length 21;
Best Local Similarity 70.0%; Pred. No. 1.8e+03;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 KXINLADRM 14
| | | | |
Db 5 KXXNLADPKL 14

RESULT 15

Q9X3K2 PRELIMINARY; PRT; 21 AA.
AC Q9X3K2;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE Cytochrome b6/f complex subunit IV (Fragment).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream."
RL Limnol. Oceanog. 43:1615-1630(1998).
DR EMBL; AF070193; AAD23234.1; -.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2317 MW; CC807B66F023D54F CRC64;

Query Match 23.3%; Score 30; DB 2; Length 21;
Best Local Similarity 63.6%; Pred. No. 1.8e+03;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PKKINLADRM 14
| | | | |
Db 4 PKKPDADPKL 14

THIS PAGE BLANK (USPTO)